(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
13 March 2003 (13.03.2003)

PCT

(10) International Publication Number
WO 03/020877 A2

(51) International Patent Classification[7]: C12N

(21) International Application Number: PCT/US02/21841

(22) International Filing Date: 15 August 2002 (15.08.2002)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/316,729 1 September 2001 (01.09.2001) US

(71) Applicant *(for all designated States except US)*: **PHARMACIA CORPORATION** [US/US]; c/o Monsanto Company, Nancy Huelskamp E2NA, 800 N. Lindbergh Blvd., St. Louis, MO 63167 (US).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **DIDION, Bradley** [US/US]; Bradley A. Didion, 220 Hickory Hollow, Washington, MO 63090 (US). **BASHKIN, James** [US/US]; James K. Bashkin, 7739 Stanford Avenue, St. Louis, MO 63130 (US). **WOODARD, Scott** [US/US]; Scott S. Woodard, 916 Pheasant Woods Drive, Manchester, MO 63021 (US).

(74) Agent: **BOND, Gary**; Gary M. Bond, 800 North Lindbergh, E2NA, St. Louis, MO 63167 (US).

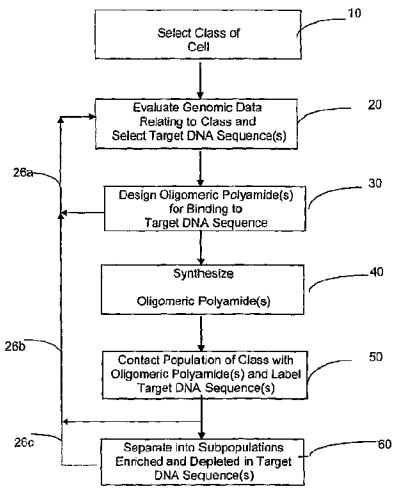
(81) Designated States *(national)*: AU, BR, BY, CA, CN, CO, IL, IN, JP, KR, MX, NZ, PL, RU, UA, US, ZA.

(84) Designated States *(regional)*: European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, SK, TR).

Published:
— *without international search report and to be republished upon receipt of that report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: STAINING AND SORTING GENOMES AND CHROMOSOMES USING SEQUENCE-SPECIFIC POLYAMIDES



(57) Abstract: Genomic DNA of reproductive cells, including sperm and other cells important in reproduction, are stained and optionally sorted by targeting oligomeric polyamides capable of fluorescence to target DNA sequences and then sorting the cells, for example, by fluorescence activated cell sorting or other cell separation techniques.

WO 03/020877 A2

STAINING AND SORTING GENOMES AND CHROMOSOMES USING SEQUENCE-SPECIFIC POLYAMIDES

FIELD OF THE INVENTION

The invention relates to sorting chromosomes or haploid genomes or certain diploid genomes by chromosomal DNA (deoxyribonucleic acid) characteristics to produce subpopulations enriched or depleted in a chromosomal DNA characteristic of interest. In another aspect, the invention relates to sorting chromosomes of a wide variety of cell types to produce subpopulations enriched in one or another chromosomal DNA characteristic. In a specific aspect, the invention relates to sorting haploid reproductive cell genomes in the form of nuclei or living or nonliving sperm, oocytes or ova or certain diploid cells such as early-stage embryos and embryonic stem cells by sex or other chromosomal DNA characteristic to produce subpopulations enriched for a particular sex determinant or for other particular chromosomal DNA characteristics.

BACKGROUND OF THE INVENTION

Artificial insemination is widely used in animal husbandry, for example, with economically important mammals such as cattle, pigs, horses, sheep, goats and other mammals and with economically important avian species such as turkeys, chickens, ducks, geese and other avians. Likewise, embryo transfer technology also has applicability and increasing application in many of these species where the value of individual offspring is sufficiently high. Both of these techniques also have human applications.

It is frequently desired to produce offspring of a predetermined sex or sex ratio, for example, female bovines for milk production or breeding, bovine males and porcine females for meat production, female poultry for egg laying and male poultry for meat production, and the like. The simplest and most economically feasible way preferentially to produce offspring of a predetermined sex or sex ratio would be a high-throughput system for producing gender-enriched semen (GES) that could then be used for artificial insemination. Another desirable technique would be to select embryos by sex to produce gender-selected embryos (GSE) and then to use the GSE for implantation for preferential production of the desired sex of offspring.

Although there are reports that sperm may be distinguishable based on sex-specific surface antigens, it is generally considered that sperm nearly completely or perhaps completely lack any phenotypic sex-specific character. As a result, efforts for producing GES in mammalian species currently rely mostly on techniques responsive to the quantitatively different levels of DNA in male and female sperm. Since in mammalian sperm, for example, a male-determining chromosome set typically has 3 to 5% total DNA less than a female-determining chromosome set, this difference has been used to separate mammalian sperm into GES using a fluorochrome that permeates the cell membranes and nonspecifically binds to the DNA without unacceptably damaging the viability of the sperm. The labeled sperm can then be sorted, for example, using cell cytometry based on the resulting quantitative difference in fluorescence between male and female sperm to produce GES. Exemplary of patent literature in this area are: WO 84/01265, Johnson et al., US 5,135,715, and Rens et al., US 5,985,216, which

are hereby incorporated by reference for techniques of sperm handling and staining and of flow cytometry applied to sperm separations.

Turning to more general aspects of reproduction, it would clearly be highly desirable to be able to select living reproductive cells based on the presence or absence of specific chromosomal DNA characteristics, for example, specific chromosomes, genes (including transgenically-introduced genes), traits, loci, and the like. The resulting populations could then be used to produce mature individuals having or lacking that particular chromosomal DNA characteristic.

In particular, it would be highly desirable to select genomes including nuclei or chromosomes of reproductive cells, such as sperm, oocytes, ova and the like, and certain diploid cells such as early-stage embryos (up to and including, for example, about the sixteen or thirty-two cell morula stage) which could be used for fertilization or implantation or otherwise for development of offspring based on the presence or absence of specific chromosomal DNA characteristics, including numbers and types of chromosomes, genetic markers, trait loci, and alleles, as well as other genetic markers such as SNPs (single nucleotide polymorphisms). The term "reproductive cells" is used in its usual and accepted sense to mean any of the haploid germ cells, usually the gametes, and the immediate predecessors from which they arise by cell division. Cf. Rieger et al., Glossary of Genetics, 5th Ed. (1991). As indicated the early-stage embryos can include zygotes and embryos up through about the morula stage of development corresponding to about the 16 or 32 cell stage of cell division but not extending beyond the blastocyst stage.

In the case of animals used for food production, it would be particularly desirable to select haploid genomes, e.g., sperm, oocytes, ova and the like based on the presence or absence of genetic markers for phenotypic traits which would be expressed in mature animals produced from such cells. A genetic marker may or may not be the location of a particular gene or allele associated with a trait of interest since it is sufficient if the genetic marker is representative of occurrence or nonoccurrence of the trait. Markers for traits might include markers for sex or other polygenic traits or for specific trait loci or alleles. Markers for polygenic traits might include markers for "QTL" or "Quantitative Trait Loci" and "ETL" or "Economic Trait Loci" which are terms used to refer to loci that may be scattered around a genome responsible for polygenic inheritance of particular characteristics.

Turning now to other areas where use of the invention to be hereinafter described would be useful, it is desirable to isolate selected nuclei of haploid cells and use them for nuclear transplantation and other developing technologies, and chromosomes of a wide variety of cells and use them for various commercial purposes including identifying suitable selectively staining molecules for sequences therein, and cells and cell clusters and to maintain them in cell suspension culture and to regenerate tissues, organs, and mature individuals from such cells. The stained sperm nuclei in accordance with the invention as hereinafter described are also highly advantageous and useful for tuning FACS instruments and use as controls while separating living sperm because the characteristic paddle-shaped nuclei and staining characteristics of sperm nuclei provide the best model for such uses. It would therefore be desirable to

provide a new technology for selecting chromosomes or reproductive cell genomes or cells for the presence or absence of particular genetic markers, trait loci or alleles. All of these and other uses and applications of the invention in its various aspects will be apparent to those skilled in the art from the description and the claims hereinafter set forth.

SUMMARY OF THE INVENTION

The invention relates to producing populations of chromosomes from a wide variety of cells or reproductive cell genomes corresponding to or comprising living cells characterized by specific target DNA sequence(s). According to the invention, a population of chromosomes or of reproductive cell or early embryonic stage genomes is provided which has been contacted with target DNA sequence specific oligomeric polyamide(s) under conditions effective for causing binding to target DNA zone(s) of dsDNA therein. Then chromosomes or genomes in the population are allocated into one or more subpopulations based on differentiating chromosomes or genomes or cells based on the extent of oligomeric polyamide(s) binding to the target DNA zone(s) of dsDNA.

In some of its various aspects, the invention relates to method, means and compositions for identifying and sorting reproductive cell genomes in the form of nuclei, killed or apoptotic or living cells by chromosomal DNA characteristics. In a specific aspect, the invention relates to sorting genomes comprising or corresponding to haploid reproductive cells by genetic markers indicative of genotypic or phenotypic traits such as sex or other traits of interest. The sorted genomes from or comprising single cells can then be used in various ways known to those skilled in the art in human

or nonhuman animal reproduction. Such cells can include, without limitation, haploid cells such as gametes, oocytes, sperm or ova, diploid cells such as zygotes, embryos, embryonic stem cells ("ESCs"), and other totipotent cells and multicellular clusters that can be used to produce or regenerate cells, tissues, organs or individuals. The identification and sorting can be done based on the type or number of sex or other chromosomes present or based on specific genetic markers, loci or alleles present or absent in the cellular chromosomal DNA.

According to another aspect of the invention, genomes of or the reproductive cells themselves (gametes, sperm, oocytes, ova), of humans or of animals, especially economically important animals such as cattle, pigs, sheep, horses, turkeys, chickens and the like, can be sorted according to sex to produce, for example, sorted genomes of various kinds or gender enriched sperm (GES), gender selected embryos (GSE) and the like for use to produce preferentially male or female offspring. According to another aspect of the invention, genomic DNA of or such reproductive cells themselves are sorted according to target DNA sequences selected to serve as genetic markers in chromosomal DNA for traits, loci or alleles to produce marker enriched, trait enriched, loci enriched, or allele enriched genomes of or reproductive cells themselves for use to produce offspring carrying such genetic markers, traits, loci or alleles.

According to another aspect of the invention, genomic DNA of or such cells themselves are sorted by a method comprising using oligomeric polyamides to label an extragenic target DNA sequence which is associated with and serves as a genetic marker for a chromosomal DNA characteristic of interest, selecting genomic DNA of or

cells themselves from a population of labeled genomic DNA or cells based on the presence or absence of the label, and producing subpopulations of genomic DNA or cells enriched or depleted in the chromosomal DNA characteristic of interest.

In a particular aspect, a method in accordance with the invention involves selecting genomic DNA or single cells or embryos based on presence or absence of specific selected target DNA sequences. The method comprises the steps of providing a population of genomic DNA or of killed, apoptotic or living cells or organisms which may contain the target DNA sequence, the specific target DNA sequence being labeled with one or more target DNA specific oligomeric polyamides, preferentially selecting the genomic DNA or the cells or organisms themselves based on the label, and producing subpopulations of genomic DNA or cells enriched or depleted in the specific target DNA sequence.

In another particular aspect, the invention relates to contacting genomic DNA of living cells with cell-permeant sequence specific oligomeric polyamides or cell-permeant formulations thereof capable of preferentially binding to specific target DNA sequences and separating the cells or clusters according to chromosomal DNA characteristics indicative of sex, trait loci or alleles.

In another particular aspect, the invention relates to the design and use of new cell-permeant sequence specific oligomeric polyamides specifically targeted to specific target DNA sequences which are effective for labeling the specific target DNA sequences in living cells and for enabling separation of cells into subpopulations based on the presence or absence or the label.

The person skilled in the art will readily realize that the products of the invention in its various aspects such as without limitation sorted genomic DNA, chromosomes, chromosomal material, killed and apoptotic and living reproductive cells and the like can be used, for example, in diagnosis and by correlation for identification of donor and recipient cells, in connection with a wide variety of reproductive and assisted reproductive technologies, including without limitation intra uterine insemination, artificial insemination (AI), *in vitro* fertilization (IVF), embryo production, embryo micromanipulation and transfer, nuclear transfer or transnucleation technologies, techniques for modification of genetic character of genetic cells, intracytoplasmic sperm injection (ICSI), chromosomal fragment identification and removal, cytoplasmic transfer preimplantation genetic diagnosis (PGD), and the like.

The invention will be further described in detail and in terms of specific preferred embodiments; however, other uses, applications and embodiments will be apparent to, or readily developed without undue experimentation by, those skilled in the art from the following detailed description and the examples.

SUMMARY OF ABBREVIATIONS AND CONVENTIONS

In describing DNA sequence data or search criteria for target DNA sequences, the IUB-IUPAC symbols for nucleotide nomenclature, including symbols for nucleotide ambiguity, will be used.

Symbol	Meaning	Nucleic Acid
A	A	Adenine
C	C	Cytosine

G	G	Guanine
T	T	Thymine
U	U	Uracil
M	A or C	
R	A or G	
W	A or T	
S	C or G	
Y	C or T	
K	G or T	
V	A or C or G	
H	A or C or T	
D	A or G or T	
B	C or G or T	
X	G or A or T or C	
N	G or A or T or C	

Cf. Cornish-Bowden (1985) Nucl. Acids Res. 13: 3021-3030

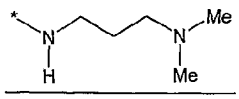
The following abbreviations may be used for oligomeric polyamide structural elements: Im = Imidazole; Py = Pyrrole; Hp = Hydroxypyrrole; Pyr = 2-pyridine; β = β-alanine; γ = γ-aminobutyrate; γN =(R) α-amino-γ-aminobutyrate; Ac = acetyl; Fo = formyl; Et = ethyl; Pr = propyl; Dp = Dimethylaminopropyl; Ed = ethylene diamine; Ta ("Triamine") = $NHCH_2CH_2CH_2N(Me)CH_2CH_2CH_2NH$; FDA = fluorescein-5-thiourea diacetate; FITC = fluorescein-5-thiourea; TRITC = tetramethylrhodamine-5-thiourea. Abbreviations for other compounds will be defined when introduced.

The abbreviations and structures are shown in the Table below:

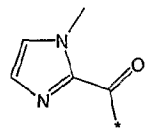
Table: Abbreviations and Structures

(* denotes the attachment bond)

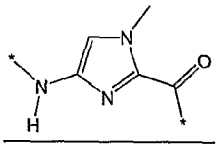
Dp ((3-dimethylaminopropyl)amine)



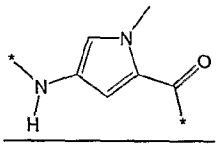
[d]Im- or Im- (1-methyl-2carboxy-imidazole, when Im is the terminal group, it is the desamino imidazole below, even if the superscript [d] is not present)



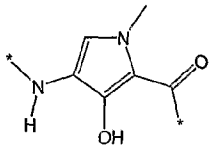
-Im- (4-amino-1-methyl-2-carboxy-imidazole, when Im is not the terminal group, it has an amino substituent)



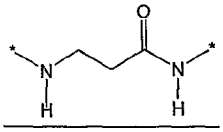
-Py- (4-amino-1-methyl-2-carboxy-pyrrole)



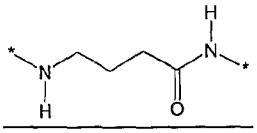
-Hp- (4-amino-3-hydroxy-1-methyl-2-carboxy-pyrrole)



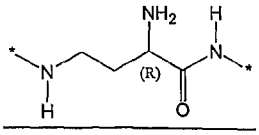
- - (beta-alanine)



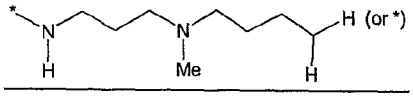
- - (gamma-aminobutyric acid)



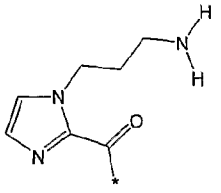
- $_{N}$- (alpha(R)-amino-gamma-aminobutyric acid, substituents on the -amino nitrogen are written - $_{N(R)}$-)



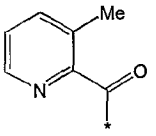
-Ta or -Ta- (triamine, 3-(3-aminopyropyl(methyl)amino)propylamine)



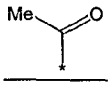
It3- (substituents on the NH2 nitrogen would be written as It3(R)-)



Pyr- (3-methylpicolinic acid)

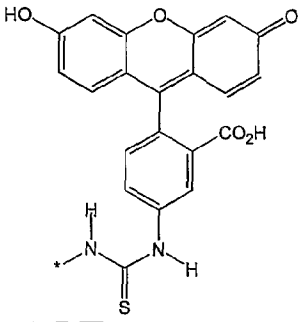


Ac- (acetyl)

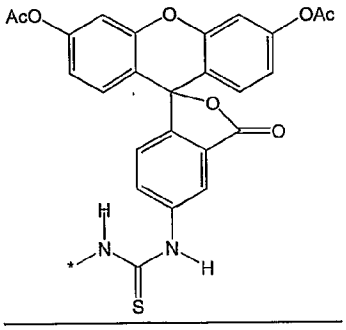


Fo- (formyl)

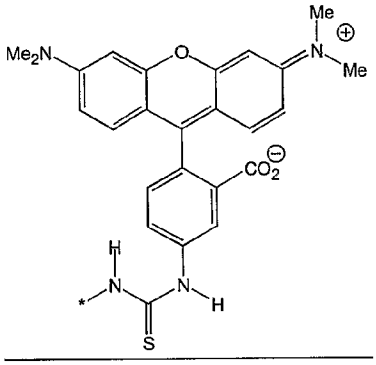
-FITC (Fluorescein isothiocyanate)



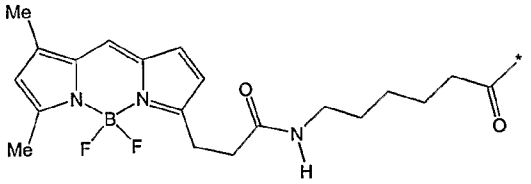
-FDA (Fluorescein isothiocyanate diacetate)



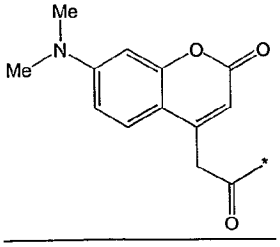
-TRITC (Tetramethylrhodamine isothiocyanate)



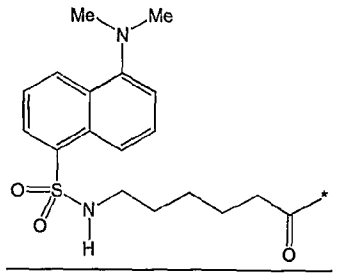
-BOFLX or -BODIPY-FL-X



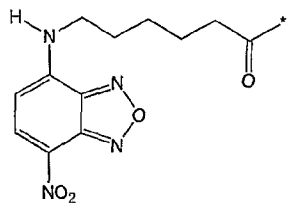
-DMACA



-dansyl-X



-NBD-X



The following conventions may be used for oligomeric polyamides: X_1X_2. . .X_n, where n is from 3 to about 30, where each X is a carboxamide independently selected from terminal linkage groups, heterocyclic carboxamides, and intercalary linkage groups (including spacer linkage groups, hairpin linkage groups, and tandem linkage groups), and where each X is connected to an adjacent X by amide linkage(s). The phrase "linear X_1X_2. . .X_n" may be used to illustrate a linear strand or linear portion of a molecule, that is, a molecule or portion (subchain) of a molecule not including a hairpin linkage group or a tandem linkage group.

BRIEF DESCRIPTION OF THE DRAWINGS

Turning now to the Drawings,

FIGURE 1 illustrates by a block diagram methods according to the invention.

FIGURE 1A illustrates by a block diagram an alternative method according to the invention.

FIGURE 2 illustrates schematically flow cytometric means and methods for separating stained chromosomes or genomes of haploid reproductive cells such as

chromosomes, nuclei and living or nonliving cells or embryos according to the invention.

DETAILED DESCRIPTION OF THE INVENTION

For convenience, the Detailed Description is subdivided into topics as follows:

1. Overview
2. Genomic DNA and Single Cells
3. Selection of Target DNA Sequence
4. Providing Oligomeric Polyamides for Target DNA Sequences
    a. General
    b. Building Blocks
    c. Labeling and Fluorescence
    d. Pairing Rules
    e. Molecular Conformations
    f. Synthesis
4. Providing genomic DNA and Cells
5. Labeling genomic DNA and Cells Separating genomic DNA such as Chromosomes, Nuclei and Cells
    a. Chromosomes
    b. Nuclei and Cells
7. Iteration of Steps
8. Examples

1. Overview

Turning now to FIGURE 1, FIGURE 1 is a block diagram illustrating various aspects of the methods in accordance with the invention. Specifically, reference numeral 10 indicates a step (described in more detail in Section 2 below) of selecting a class of cells of interest as a source of genomic DNA including the cells themselves. For example, the class may be selected by species and by cell type, such as reproductive cells like gametes, including sperm, oocytes, and ova, as well as certain diploid cells such as zygotes, embryos, embryonic stem

cells, of humans or of animals important in animal husbandry or as companion animals, such as cattle, swine, horses, chickens, and the like.

Reference numeral 20 indicates, after class selection, a step (described in more detail in Section 3 below) of evaluating available genomic data relating to the selected class to select target DNA sequences. In some cases, all that may be known is the number or types of chromosomes. In most cases, however, there will be known and increasingly become available more detailed sequence information and genetic map and physical map information relating to the class which may be searched and otherwise used for selecting appropriate genic or extragenic target DNA sequences associated with a chromosomal DNA characteristic of interest.

Reference numeral 30 indicates a step (described in more detail in Section 4 below) of designing oligomeric polyamides for optionally cell permeation and binding to the selected target DNA sequences with high affinity and specificity. Design of oligomeric polyamides for specific target DNA sequences is often an iterative process involving repeated applications of steps 20 and 30 as indicated by reference numeral 26a, or even of steps 20, 30, 40 and 50 (described below) as indicated by reference numeral 26b.

Reference numeral 40 indicates a step (described in more detail in Section 4 below) of synthesizing the designed oligomeric polyamides.

Reference numeral 50 indicates a step (described in more detail in Section 5 below) of bringing the synthesized oligomeric polyamide(s) into contact with genomic DNA of the selected class of cells under conditions effective for causing at least some of the target DNA sequences to become labeled.

Reference numeral 60 indicates a step (described in more detail in Section 6 below) of separating or sorting the labeled genomic DNA of the cells into subpopulations enriched or depleted in the target DNA sequences and in the chromosomal DNA characteristic of interest with which the target sequence is associated.

Reference numeral 26b indicates a step (described in more detail in Section 7 below) of repeating steps 20, 30, 40, 50, and possibly 60, to develop effective or optimized systems for practicing the different aspects of the invention.

Referring now to FIGURE 1A, FIGURE 1A illustrates alternative embodiments of the method according to the invention in which chromosomal genomic DNA is isolated and sorted using oligomeric polyamides. Specifically, after selecting the class of cell of interest (reference numeral 10'), the chromosomes are isolated and separated as indicated by reference numeral 15 from the class of interest or from a suitable culture of cells of the same species of animal. This can be conducted, for example, by flow cytometry as discussed below in Section 6.a. in more detail. Thereafter the isolated and separated chromosomes can be contacted with available or specially constructed polyamide(s) (reference numeral 25) to select polyamides that preferentially bind to chromosome(s) of interest. Then the selected class of cells can be contacted with the selected polyamides (reference numeral 50') and separated into subpopulations (reference numeral 60') enriched or depleted relative to each other in the selected polyamide(s) thereby producing subpopulations enriched or depleted in the chromosome or chromosomes of interest. The same approach can of course also be used to selected target DNA sequences of greater specificity.

2. Single Cells and Nuclei

According to the invention, populations of reproductive cell genomic DNA or living or nonliving cells or isolated nuclei thereof are provided labeled with oligomeric polyamides, especially oligomeric polyamides carrying fluorophores (polyamide-dye conjugates) or other labels, specific for specific target DNA sequences indicative of chromosomal DNA characteristics of interest. The classes of cells can be any class of prokaryotic cells such as bacteria or of eukaryotic cells, including animals, plants and eukaryotic fungi, for which such target DNA sequences can be identified and which are capable of separation by such techniques, especially those which are capable of being processed by flow cytometry or by micromanipulation.

While it has been well known that some DNA fluorophores are capable of passing through the sperm plasma membrane and binding quantitatively to the DNA (e.g., Hoechst 33342 excited by UV light) allowing separation into X and Y chromosome bearing populations, such has not been demonstrated for polyamide dye conjugates. Further, the Hoechst fluorophore requires ultraviolet light for excitation with the possibility of causing ultraviolet damage whereas the use of polyamide dye conjugates for separation of sperm into X and Y populations can be performed using wavelengths of light that are distanced from UV excitation (i.e., in the visible light range) or can of course also be performed using ultraviolet excitation.

Preferably, the genomic DNA or cells are selected from reproductive cells of animals including gametes (sperm, oocytes and ova), or even zygotes and early-stage embryos and embryonic stem cells, and the like, capable of being maintained in cell suspension culture and of being processed, for example, using flow cytometry or

micromanipulation to produce subpopulations enriched or depleted in the target DNA sequences of interest.

There will usually be a difference of uptake of a particular polyamide or polyamide-dye conjugate depending on the particular cell type. For example, as is well known, cells have cell surface characteristics including cell surface receptors of several kinds (signaling receptors, adhesion receptors, transport receptors, and the like) and further may differ significantly in membrane structure and composition. For these or other reasons, there are observed differences in uptake of polyamides and polyamide-dye conjugates depending on the size and structure of the polyamide and the type and character of the conjugated dye for particular types of cells. Consequently, it will be usually necessary to evaluate whether the cells of interest effectively take up particular polyamide or polyamide-dye conjugates, and if they do not, to select different polyamides or polyamide dye conjugates which are effectively taken up by the cell under the conditions of usage. In the alternative, the cells can be treated as discussed below in more detail so as to facilitate uptake of the particular polyamide or polyamide-dye conjugate, for example, by treatment with particular solvents such as DMSO and the like, or by use of liposomes or electroporation or cell-permeation-enhancing solutions or other techniques for facilitating uptake already known to those skilled in the art, for example, in connection with introducing DNA into cells.

In accordance with a preferred aspect of the invention, the living cells may comprise human or animal sperm to be used for artificial insemination or in vitro fertilization. Alternatively, the cells may be human or

animal oocytes or ova or even zygotes or early-stage embryos to be used for embryo transfer.

In particular aspects, the living or nonliving cells are selected from sperm, ova, oocytes, etc. zygotes and embryos of mammals (not excluding humans) such as cattle, pigs, horses, sheep, as well as others, or of avian species such as turkeys, chickens, ducks, geese and the like, where there is a significant difference in total chromosomal DNA depending on whether the X or Y chromosome (W or Z in avians) is present or in what numbers. Thus, for example, it is possible to distinguish mammalian X-bearing and Y-bearing sperm based on difference in total chromosomal DNA present, and likewise to distinguish avian W-bearing and Z-bearing sperm. Alternatively, it is possible to distinguish such sperm based on DNA sequences specific to the different sex chromosomes, that is, based on Y-specific, X-specific, W-specific and Z-specific sequences.

3. Selection of DNA Target Sequences.

According to the invention, target DNA sequences are selected and used for design and synthesis of oligomeric polyamides. As used herein, the terms "target DNA sequence" and "target DNA region" may be used to refer to the DNA sequence selected for polyamide binding or to the sequence to which the polyamide binds. In connection with the design of polyamides, the length of the target DNA sequence or region is considered to be the entire DNA sequence which is selectively bound by the polyamide, i.e., all of the bases or base pairs which contribute to the selectivity and affinity of the particular polyamide. Typically, this includes the heterocyclic binding pairs and also the linkers, protonatable amines and other building blocks of the particular polyamide.

The target DNA sequences or regions may be selected based on theoretical considerations such as expected frequency or degree of occurrence of specific DNA sequences in genomes based on considerations of DNA length, GC:AT ratios, and the like, or selected specifically by searching available DNA sequences of the selected class of cells for suitable targets. Such available sequence data may be gene-related DNA sequences referred to herein by the term "genic". The term "genic" is thus used to refer to sequences, regions, and the like which are closely and functionally associated with expression or control of expression of the coding sequences (exons) of genes, whether transcribed with the gene such as 5'-UTRs, 3'-UTRs, leaders, introns, etc. or serving a regulatory function such as promoters, TATA boxes, enhancers, *cis*-acting regulatory sequences, etc. In some instances, for example, in targeting specific alleles or where genic sequences are known but suitable extragenic genetic markers cannot be identified, it will be preferred to use genic sequences for selection of target DNA sequences.

In accordance with an aspect of the invention, however, it is frequently preferred to search extragenic DNA sequences for target DNA sequences associated with a chromosomal DNA characteristic of interest. As is well known, in eukaryotic species, depending on the species, large quantities of chromosomal DNA may be extragenic DNA of little or no known function. Such extragenic DNA may be DNA sequences occurring in moderately or highly repetitive repeats or dispersed or clustered repeats characteristic of genomes, including SINEs (short interspersed nuclear elements), LINEs (long interspersed nuclear elements), satellite DNA, minisatellite DNA, and microsatellite DNA. Such extragenic DNA often provides a

highly enriched source of potential target DNA sequences for facilitating oligomeric polyamide design. By selecting extragenic DNA sequences as target DNA sequences one obtains a much larger set of DNA data which can be searched for appropriate target DNA sequences and moreover can select sequences associated with but not identical with or related to the functioning of genes of interest, thereby minimizing likelihood of any interference with gene function by oligomeric polyamide labeling.

In accordance with the invention, a DNA target sequence is selected and a fluorophore, which is an oligomeric polyamide or is carried by one or a combination of oligomeric polyamides, is provided (described below) which preferentially binds adjacent the target DNA sequence.

Generally speaking, there are multiple different approaches which may be followed in selecting target DNA sequences of interest: One approach is to select target DNA sequences which occur or are expected to occur approximately randomly in the genome but where the physical amount or quantity of chromosomal material present in cells or multicellular clusters differs as a function of the chromosomal DNA characteristic of interest thereby permitting sorting based on differences in total labeling between cells or clusters. This is one approach followed, for example, in distinguishing Y-bearing from X-bearing sperm, since the difference in total chromosomal material between X and Y sperm permits separation. This approach can also be used in separating aneuploid cells or polyploid cells or cells with gross chromosomal structure deviations, including cells containing chromosomal fragments, and the like from normal cells.

Another approach is to select target DNA sequences that occur at a higher or lower frequency as a function of the chromosomal DNA characteristic of interest. For example, extragenic DNA characterized by moderately or highly repetitive sequences is known to be characteristic of particular chromosomes. Some of these repetitive sequences are unique to or highly localized on the sex chromosomes. Consequently, this approach can also be used in separating Y-bearing from X-bearing sperm, since it is known that mammalian Y chromosomes contain multiple unique noncoding 'nonsense' repeats which can be searched for shorter sequences occurring in X or Y chromosomes at a higher, or at a lower, frequency or total amount or quantity than would be expected in the genome as a whole. Use of these shorter target DNA sequences permits separation of cells containing one or more X- or Y-chromosomes from cells not containing or only containing one X- or Y-chromosome. This approach can also be used in connection with target DNA sequences, especially repetitive sequences, which serve as markers for specific traits, genes or alleles or other chromosomal DNA characteristics.

Another approach in determining a target DNA sequence is to select a DNA sequence known to occur in genomic DNA but which is of sufficient length to make it likely to occur uniquely or only in low copy number in the genome. For example, unique or low copy number extragenic DNA sequences are known to constitute a large percentage of genomic DNA in many species, including humans and animals used by humans for meat, food products, work or companionship. In addition, genic target DNA sequences (genes and gene-related sequences such as associated regulatory sites, introns, etc.) can be targeted using this approach. As before, this

approach can also be used in connection with separating Y-bearing from X-bearing sperm, but can also be used for other applications of the invention.

Another approach is empirical and involves creation of combinatorial libraries of oligomeric polyamides that are then tested against cell populations or chromosomes or DNA sequences of interest to identify polyamides having a higher, or lower, affinity for certain of the cells of interest. The oligomeric polyamide, as will be appreciated from the description below, then defines the target DNA sequence.

Preferably, the target DNA sequence is selected to have a higher or lower frequency or quantity of occurrence in the genomes containing the chromosomal DNA of interest than would be expected on a random occurrence basis.

To further illustrate this aspect of the invention, reference may be had to FIGURE 1A that illustrates an alternative method in accordance with the invention in which chromosomes are separated and isolated from cells of the species of interest and are used to select polyamide(s) that preferentially bind to those chromosome(s). The polyamides in these cases therefore determine the target DNA sequences.

Referring now in more detail to FIGURE 1, after a class of cells of interest is selected, chromosomes characteristic of the cells can be isolated and separated using any technique known to those skilled in the art, for example, by flow cytometric analysis of chromosomes. This methodology is further illustrated below in Section 6.a.

Broadly, the DNA target sequence can be in the range of 2 to about 30 bases in length, preferably in the range of about 4 to about 22 bases in length, more preferably

in the range of about 8 to about 16 bases in length, or even in the range of about 16 to about 22 bases in length. Those skilled in the art will appreciate that base sequences in the range of about 16 to 22 bases in length are extremely unlikely to occur on a random basis and therefore provide the highest specificity of sequence. Sequences in the range of 8 to 16 bases may empirically occur at a higher or lower than random frequency in particular DNA sequences of interest than in the genome as a whole and therefore be suitable as target DNA sequences. Sequences in the range of 2, 3, 4 up to 8 bases, albeit expected to occur at a fairly high frequency in the genome as a whole, can still be used to distinguish cells containing chromosomes which differ significantly in total DNA from other chromosomes, such as the X, Y, W, Z chromosomes, or aneuploid cells, or deviant chromosomes occurring as a result of breakage and recombination. Those skilled in the art will appreciate that the above ranges are particularly suited to organisms or species characterized by large genomes, and that smaller sequences may achieve similar results in connection with smaller genomes, such as, e.g., the yeast genome.

According to an aspect of the invention in which at least two polyamides, for example, one bearing an energy donor and the other bearing an acceptor fluorophore, or alternatively one bearing an donor fluorophore and the other bearing an acceptor and quencher, are caused to bind adjacent each other in the minor groove of dsDNA at an effective distance between the two effective for permitting energy transfer and fluorescence or quenching of fluorescence, the target DNA sequence for each polyamide can be shorter than would be required for a single polyamide having the same target specificity.

This is because the required functional cooperation between the two polyamides in effect causes the two shorter polyamides to have about the same specificity as would a single polyamide of length corresponding to the sum of the two.

In such cases, oligomeric polyamides can be designed to bind consecutively and adjacently but not necessarily contiguously along the target DNA sequence. Thus, it will be possible to target a longer more specific target DNA sequence while using shorter polyamides each of which is specific for only a part of the total target DNA sequence and the shorter polyamides may be easier to synthesize and more likely to permeate the cells without unduly interfering with cell viability.

Turning now more specifically to selecting DNA target sequences for distinguishing sperm based on sex, target DNA sequences can be selected which occur randomly throughout the entire genome (nonspecific or randomly distributed or non-unique target approach) or a target sequence can be selected which is specific to or non-randomly distributed or highly localized on the particular sex chromosome or on the other sex chromosome of interest (specific or non-randomly distributed or unique target approach). Where the nonspecific target approach is taken, it is preferred to select a target DNA sequence in the range of 2 to 8 base pairs thereby assuring random occurrence in the genome. If desired, the selected target sequence can be tested against publicly available DNA sequences of the species of interest to confirm randomness of occurrence. Where the specific DNA target approach is taken, it is preferred to select target DNA sequences in the range of 8 to 22 bases in length to assure nonrandom preferential occurrence in the chromosome to be labeled.

Turning now more specifically to separating genomic DNA of haploid reproductive cells such as sperm and oocytes, ova, or embryos by sex, there are known sequences which are specific to the sex chromosomes of many if not most species. These sequences include both gene and gene-related and extragenic DNA sequences. For example, see WO 88/01300, WO 89/02440, WO 89/07154, WO 90/15155, WO 91/00365, WO 92/00375, WO 92/06215, WO 97/31012, WO 99/33956, US 5,215,884, US 5,759,772, CA 2113957, EP 546,762, FR 2,635,116, JP 04-360687, JP 05-076367, JP 08-205895 and in numerous papers including McGraw et al., 16 Nucleic Acids Research 10389 (1988); Mileham et al., 16 Nucleic Acids Research 11842 (1998); Matthews et al., 13 Genomics 1267-1273 (1992); Cotinot et al., 10 Genomics 646-653 (1991); Miller et al., 21 Animal Genetics 77-82 (1990); Matthews, "Repeated DNA Sequences from the Bovine Y-Chromosome, Ph.D. Thesis, Australian National University (1990); Matthews et al., Cytogenet. Cell Genet. 56:40-44 (1991); Kudo et al., J. Reprod. Dev. 39:55-63(1993). All of these references are incorporated herein by reference for purposes of showing DNA sequences that can be searched for target DNA sequences for use in connection with various aspects of the invention.

In accordance with a preferred aspect of the invention the target DNA sequence are selected from these highly repetitious localized regions.

3. Oligomeric Polyamides Binding to Target DNA Sequences

a. General

In accordance with the invention, oligomeric polyamides adapted to lie at least partly in the minor groove of dsDNA and comprising predominantly heterocyclic monomers capable of preferential selectivity for specific DNA bases or base pairs along the minor groove and high

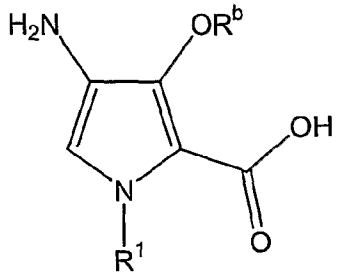
affinity for the target DNA sequence as a whole are provided. The term "oligomeric polyamides" refers to oligomers comprising one or more oligomeric linear moieties, each linear moiety comprising from 3 to 30 monomeric units connected by amide linkages (-CONH-) along the molecular chain. The phrase "adapted to lie at least partly in the minor groove of dsDNA" is used to refer to the mode of binding of the oligomeric polyamides described herein. The phrase "comprising predominantly heterocyclic monomers capable of preferential selectivity and affinity for specific DNA bases or base pairs along the minor groove" is used to refer to the aspect, described below, that the oligomers predominantly comprise heterocyclic monomeric units which specifically confer a greater specificity or affinity or both for one or more, but not all, of the DNA bases lying along the minor groove in the target DNA region. Taken as a whole, the phrase indicates molecules adapted to minor groove binding and having high target DNA sequence affinity and specificity.

Oligomeric polyamides adapted to lie at least partly in the minor groove of dsDNA and comprising predominantly heterocyclic monomers capable of preferential affinity for specific DNA bases or base pairs along the minor groove are known and have been the subject of extensive recent activity in both the scientific and patent literature. Consequently, the person skilled in the art is capable of designing, constructing and using these molecules in the manner and to the extent described in the relevant literature.

The subject compounds can be prepared using either liquid phase or solid phase methods known to those skilled in the art. For liquid phase synthesis, see, e.g., Xiao et al., "A Convenient Method for the Synthesis

of DNA-Recognizing Polyamides in Solution," J. Org. Chem. ACS ASAP. CODEN: JOCEAH ISSN:0022-3263. AN 2000:537885 CAPLUSJ; see also Boger et al., "Total Synthesis of Distamycin A and 2640 Analogues: A solution-phase combinatorial approach to the discovery of new, bioactive DNA binding agents and development of a rapid, high-throughput screen for determining relative DNA binding affinity or DNA binding sequence selectivity," Am. Chem. Soc.122, 6382 (2000). For solid phase synthesis, see, e.g., Baird and Dervan, J. Amer. Chem. Soc. 1996, 118, 6141, US 5, 998,140 and US 6,090,947. More generally, the following patents, patent applications and literature are illustrative of the information available to those skilled in the art and the contents of these documents and those hereinabove cited are incorporated herein by reference: US 5,998,140; and WO Applications WO 98/50582, WO 98/50058, WO 9849142, WO 98/45284, WO 98/37087, WO 98/37067, WO 98/37066, WO 98/35702 and WO 97/30975.

Turning in more detail to the synthesis of Hp, the monomers of interest for synthesis of oligomeric polyamides include those represented by the following formula:



wherein R^1 is typically methyl, and R^b a protective group to block side reactions during the course of the polyamide synthesis. Such compounds can be prepared, for example, as described in the references above, or as described in Example XII below, or as described in Urbach et al., "Sequence Selectivity of 3-Hyroxypyrrole/Pyrrole Ring Pairings in the DNA Minor Groove," J. Am. Chem. Soc., 1999, *121*, 11621-11629, or even as described in an earlier article of Momose et al., "3-Hydroxypyrrole. I. A General Synthetic Route to 4,5-Unsubstituted Alkyl 3-Hydroxypyrrole-2-carboxylates," Chem. Pharm. Bull., 26(7), 2224-2238 (1979). These references are incorporated herein by reference, particularly in reference to showing synthetic methods for Hp other than and in addition to those shown in Example XII below.

b. Building Blocks

Oligomeric polyamides suitable for use with the invention comprise one or more linear oligomeric portions or subchains comprising 3 or more heterocycles that associate with specific nucleotide bases in the dsDNA target sequence. When the polyamide comprises two linear portions or subchains linked by a hairpin or amino or tandem linkers (described in more detail below), the side by side configuration of the subchains in the minor groove creates antiparallel binding pairs of polyamide heterocycles which associate with complementary base pairs in the dsDNA and significantly increase specificity.

Short linkage groups of various functionalities may be associated with the heterocyclic carboxamides forming the backbone of the oligomeric polyamides. The linkage groups may occur between the heterocyclic compounds (intervening linkage groups) or may occur at the ends of the oligomers (terminal linkage groups). Intervening

linkage group functionalities may include effecting a hairpin turn at appropriate points along the backbone or effecting a hairpin turn between two linear portions or subchains and linking to an additional linear portion (tandem linkage) or adjusting alignment by adding length or flexibility in a linear portion of the molecule. Frequently, such linkage groups are constructed using aliphatic or other amino acids, particularly aliphatic amino acids having a terminal amino group. Of course, when both ends of the oligomeric polyamide have polar groups, it is desirable to avoid hindrance to insure that similarly charged polar groups are not directly adjacent each other in the minor groove.

In each linear portion of a hairpin or tandem hairpin polyamide there is preferably a series of three or more consecutive heterocyclic carboxamides forming in an antiparallel fashion binding pairs with three or more consecutive heterocyclic carboxamides in another linear portion of the oligomer. A backbone constructed to provide in each linear portion of the molecule at least three and up to 9 or more consecutive heterocyclic binding pairs is preferred. As the number of consecutive heterocyclic binding pairs increases above about 6, it may be desirable to intercalate spacer linkers to adjust length or flexibility of the linear portions or subchains to assure effective binding in the target DNA zone of the minor groove.

According to a preferred aspect of the invention, the oligomeric polyamides are Im/Py polyamides, that is, polyamides whose backbones predominantly comprise the building blocks for which abbreviations are provided above in the Summary of Abbreviations and which are exemplified in the Examples set forth herein. Such polyamides are synthetic molecules comprising amino acid

derivatives of N-methylpyrrole, N-methylimidazole, and 3-hydroxy-N-methylpyrrole. The Im/Py polyamides are preferred because they are known to bind to DNA in a sequence specific manner with high affinity, because the molecular basis of specificity is known and can be altered to selected target DNA sequences, and because exogenous Im/Py polyamides can be shown to permeate living cells where they can bind DNA in vitro. Nevertheless, the polyamides and carboxamides that may be used in accordance with the invention are not limited to the Im/Py polyamides but include generally all polyamides and carboxamides capable of permeation and specific binding to target DNA sequences including those described in the references cited above. The useful polyamides and carboxamides also include the pharmaceutically acceptable salts thereof, as well as all the possible isomers and tautomers thereof, both separately and in mixture. The salts include both the salts with pharmaceutically acceptable inorganic acids such as, for example, hydrochloric, hydrobromic, nitric and sulfuric, or with organic acids such as, for example, citric, tartaric, maleic, fumaric, methanesulfonic, and ethanesulfonic.

The heterocyclic monomer units may be substituted at positions of the monomer that are directed away from the floor of the groove for any purpose. Thus, a hydrogen atom may be substituted with a substituent of interest, where the substituent will not result in steric interference with the wall of the minor groove or otherwise create repulsion. While the invention contemplates that a wide range of substituents may be used, it may be particularly desired to use substituents that cause or increase or quench fluorescence when bound adjacent a target DNA sequence.

The intercalary linkage groups between the heterocyclic groups will generally have a length of two atoms, wherein at least some of the linking groups will have NH, where the NH may hydrogen bond with an unshared pair of electrons of the nucleotides. The linking chains may be methyleneamino, carbamyl (--CONH--), ethylene, thiocarbamyl, imidinyl, and the like, particularly carbamyl and its heteroanalogs, e.g. thio and imino.

Functional linkage groups comprising aliphatic amino acids are employed, particularly γ-amino- or αγ-diamino-aliphatic amino acids, either to provide for hairpin turns between first and second portions or subchains of oligomers to provide complementation between the two sequences of heterocycles, to form cyclic compounds where oligomers are joined at both ends, or to provide for a shift in spacing or flexibility of the organic cyclic compounds in relation to the target dsDNA. For the most part, the aliphatic amino acids will have a chain as a core structure of two to six carbon atoms, usually of two to four carbon atoms, desirably having terminal amino groups, particularly glycine, β-alanine, and γ-aminobutyric acid, being unsubstituted or substituted on carbon and nitrogen, particularly carbon, although for the most part the aliphatic amino acids will be unsubstituted.

These amino acids will play specific roles. The longer chain aliphatic amino acid will serve to provide for turns in the molecule and to close the molecule to form a ring. The shorter chain aliphatic amino acids will be employed, both to provide a shift for spacing in relation to the target dsDNA, and to provide enhanced binding by being present proximal to the terminal organic cyclic group. The aliphatic amino acid may be present at one or both ends of the oligomer. Of particular interest

is glycine and alanine; for space shifting, β-alanine is preferred. Usually, a consecutive sequence of 6 heterocycles will be avoided. Generally, there will be an amino acid, particularly β-alanine, introduced in an otherwise consecutive series of six oligomer units, generally bordered by at least two, preferably at least three organic cyclic groups, particularly heterocycles.

The aliphatic chains of the aliphatic amino acids may also serve as sites of substitution, the aliphatic amino acid providing a core structure, there usually being not more than 2, more usually not more than 1, substituent. The same types of substituents that have been described for the heterocycles may also be employed here. Conveniently, the substituted aliphatic amino acid may be used in the synthesis of the oligomer, rather than modifying the amino acid after the oligomer is formed. Alternatively, a functional group may be present on the chain of the substituent, if necessary being appropriately protected during the course of the synthesis, which functional group may then be used for the subsequent modification. Desirably, such functional group could be selectively used, for synthesis of different oligomers, so as to provide for substitution at that site to produce products having unique properties associated with a particular application. With the substituent substituted at a site which does not significantly interfere with the binding in the groove, e.g. employing a single stereoisomer, properties can be imparted to the subject compounds, such as water solubility, lipophilicity, permeability, metabolic stability, non-covalent binding to a receptor, radioactivity, fluorescence, etc.

One or both termini of an oligomer, preferably one of the termini, will have a polar group substituted on an

alkyl group, where the polar group will generally be from 2 to 6, more usually 2 to 4, carbon atoms from the linkage to the remaining molecule. The polar group may be charged or uncharged, where the charge may be a result of protonation under the conditions of use. Particularly, groups capable of hydrogen bonding are preferred, such as amino, particularly tertiary-amino, hydroxyl, mercapto, and the like. Of particular interest is amino, more particularly alkylated amino, where the alkyl groups are of from 1 to 6, usually 1 to 3, more usually 1, carbon atom, and at a pH less than about eight, the amino group is positively charged, and can hydrogen bond with the dsDNA. Desirably, two positively charged polar groups would not be employed on the oligomers, where the positively charged polar groups will be in juxtaposition when complexed with the dsDNA. It is found that the presence of the two positively charged polar groups in proximity tends to reduce the binding affinity of the oligomer.

Besides the other sites present on the oligomer, either terminus of the oligomer may be adapted for detection. For example, the oligomer may be linked to labels that do not unacceptably reduce viability of the cells or permeation into the cell. Such labels can be associated with a single oligomeric polyamide, or alternatively two polyamides can be provided which are effective to cooperate for causing labeling. See Section 4.c. below.

The different molecules may be joined to the termini in a variety of ways, depending upon the available functionality (functionalities) present at the termini, such as extending the polar substituted alkyl group, e.g. having a chain of more than 6 carbon atoms, providing for a substituent at a terminus which can be reacted with the

moiety to be added, where such substituents will conventionally be amino, hydroxyl, mercapto, carboxyl, phosphate, etc., so as to form amides, both organic and inorganic, substituted amines (reductive amination), ethers, thioethers, disulfides, esters, both organic and inorganic, pyrophosphates, and the like. The molecules may be introduced as part of the synthetic scheme, displacing the oligomer from the solid support on which the oligomer is synthesized. Because the compounds of the subject invention may be used in such a variety of ways, no simple description is appropriate to the variety of moieties to which the subject oligomers may be bound, nor the specific molecular weights of the resulting products.

c. Labeling and Fluorescence

According to the invention, oligomeric polyamides are labeled with a suitably detectable label for permitting separation of cells and cell clusters based on the presence or absence of target DNA sequences therein. For various reasons, fluorescent labels or labels which quench fluorescence when binding in the minor groove of dsDNA along the target DNA zone are preferred.

We have found that some oligomeric polyamides are directly capable of fluorescence when binding in the target zone of the minor groove. Thus the oligomeric polyamides may be selected to be capable of fluorescing directly or may be adapted to fluoresce by adding a bound fluorochrome as a substituent group or by adding bound moieties which interact to fluoresce or to quench fluorescence when bound adjacent one another in the target DNA zone of dsDNA.

Alternatively, oligomeric polyamide - dye conjugates can be prepared and used in accordance with the

invention. Generally, it is preferred that the dye be functionalized and conjugated to the polyamide and that the conjugated polyamide-dye be cell and nucleus permeant. A dye whose fluorescent properties are activated intracellularly can be used if the activated dye can permeate the nucleus. If the dye has any DNA binding affinity independently of the polyamide to which it is attached, such affinity should not be detrimental to the affinity and specificity of the polyamide for the target DNA sequence.

Amine-reactive fluorophores that can be readily used to modify the oligomeric polyamides used in accordance with the invention are preferred though any suitable conjugatable fluorophore may be used. Amine-reactive dyes react with non-protonated aliphatic amine groups, including the amine terminus of polyamides as described herein, and the ε-amino of lysines and can readily be conjugated to polyamides as described herein using art-recognized techniques for amines. Uncharged or neutral polyamide-dye conjugates are particularly effective at permeating the cell and binding in the nucleus and are preferred. Suitable compounds, together with descriptions of methodologies useful for conjugating the dyes to amines, are available from Molecular Probes Inc., PO Box 22010, Eugene OR 97402-0469. Exemplary of the compounds which may be used to form such conjugates are, for example, FDA (fluorescein-5-thiourea diacetate), BODIPY FL-X, DMACA, dansyl-X, NBD-X, and the like.

According to a preferred aspect, fluorescence resonance energy transfer (FRET) systems can be conjugated to and used with the oligomeric polyamides in practicing the various aspects of the invention. FRET is a distance-dependent interaction between the electronic

excited states of two dye molecules in which excitation is transferred from a donor molecule (D) to an acceptor molecule (A) without emission of a photon.

The primary conditions for FRET include (i) that the D and A molecules must be in close proximity (typically 10-100 Å), (ii) that the absorption spectrum of A must overlap the fluorescence emission spectrum of D, and (iii) that D and A transition dipole orientations must be approximately parallel. The Förster radius (R_o) defines the distance at which energy transfer is 50% efficient (i.e., 50% of excited donors are deactivated by FRET). The magnitude of R_o is dependent on the spectral properties of D and A. In most applications, the donor and acceptor dyes are different, in which case FRET can be detected by the appearance of sensitized fluorescence of the acceptor or by quenching of donor fluorescence. D and A pairs comprising non-fluorescing donors and fluorescing acceptors or comprising fluorescing donors and quenching acceptors may be used in accordance with the invention. When D and A are the same, FRET can be detected by the resulting fluorescence depolarization.

D and A dye-pairs that can be functionalized and readily conjugated to oligomeric polyamides where the resulting conjugated polyamide-dyes are cell and nucleus permeant are preferred. A currently preferred dye pair is BODIPY-FL-X and BODIPY-TMR-X since these have carboxylic acid groups for coupling to the polyamide using activated ester technology known to those skilled in the art. Extensive compilations of D and A dyes and R_o values can be found in the literature and the person skilled in the art may select suitable dye-pairs from such compilations by following the teachings herein.

Since FRET efficiency is influenced by distance between A and D and also influenced by the angle between the dipole moments of D and A, it is preferred to select target DNA sequences to provide for D and A being at an efficient distance and appropriate dipole transition alignment for energy transfer. As applied to β form dsDNA, in which the pitch of the dsDNA helix is 34 Angstrom corresponding to 10 bp distance and with a diameter of about 20 Angstrom, this means that the target DNA sequences should be such that the distance between the A, D molecules conjugated to the polyamides is less than about 30 bp, preferably less than about 20 bp, most preferably less than about 10 bp, and in each case with the D and A molecules preferably having dipole transitions which are approximately parallel.

d. Molecular Conformation

In discussing polyamide design and target DNA binding, it will be useful to illustrate oligomeric polyamides relative to the target DNA zone of dsDNA structure. In such cases, dsDNA structure will be shown with the sense or template strand reading left to right in the 5′→3′ direction in an upper position and the complementary 3′→5 strand, if shown, in a lower position. In dealing with introns, 5′ or 3′ untranslated regions and other noncoding or nonsense regions when a contiguous or nearby coding region is known, the nearby coding region will be used to determine which strand is considered the "sense" strand. When no nearby known coding sequence is known, as may be the case with extragenic regions, nonsense repeats and the like, either of the DNA strands may be arbitrarily selected as the upper strand for illustration.

Molecular conformations for oligomeric polyamides for binding to a particular target DNA sequence can be any conformation having at least one or preferably two linear oligomeric polyamides or linear portions or subchains of a polyamide capable of lying side by side in the target DNA zone of the narrow groove. Such conformations include linear 1:1 and 2:1 complexes, the latter including both fully overlapped and slipped motifs, hairpin, cyclic, amino hairpin, H-pin, and tandem hairpins. These conformations are illustrated and described below in more detail. In general, the folded conformations are preferred and the hairpin folded conformations are most preferred. However, for particular applications, such as for using the nonspecific DNA target approach described above for sorting sperm, linear conformations may be preferred to facilitate synthesis and permeation into the cells and nuclei.

Shown below are a 1:1 complex formed between a single linear Py-Im oligomer and a target dsDNA sequence, (SEQ ID NO:1) and, and a 2:1 complex formed between two linear Py-Im oligomers and a target DNA sequence (SEQ ID NO:2).

1:1 Complex	2:1 Complex
5'-T-A-T-G-T-A-T-T-T-A-T-3'	5'-T-A-T-G-T-A-T-C-T-A-T-3'
Im-Py-Py-Py-Py-β-Dp	Im-Py-Py-Py-Py-β-Dp
	Dp-β-Py-Py-Py-Py-Im
3'-A-T-A-C-A-T-A-A-A-T-A-5'	3'-A-T-A-C-A-T-A-G-A-T-A-5'

Except for the cyclic polyamides, all polyamides have the potential to bind in a linear fashion in a manner such as illustrated. The resulting polyamide-dsDNA complexes are believed to have the polyamide

positioned approximately in the middle of the target DNA zone of the minor groove with the amides forming complementary pairs via hydrogen bonds to the adjacent DNA base pairs. However, the polyamides are unable to distinguish A/T from T/A or G/C from C/G resulting in the 1:1 binding complex being less preferred for many applications of targeting DNA sequences but simplifying design of the polyamides since it is only necessary to search a DNA sequence for appropriate targets without distinguishing A/T and T/A pairs or C/G or G/C pairs.

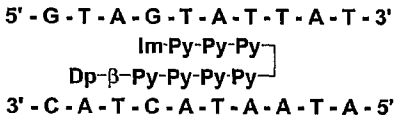
Shown below are a 1:1 complex formed between a single linear Py-Im oligomer and a target dsDNA sequence (SEQ ID NO:3), and a 2:1 complex formed between two linear Py-Im oligomers and a target dsDNA sequence (also SEQ ID NO:3).

1:1 Complex	2:1 Complex
5'-T-A-T-G-T-A-T-G-T-A-T-3'	5'-T-A-T-G-T-A-T-G-T-A-T-3'
Im-Py-Py-Py-Im-β-Dp	Im-Py-Py-Py-Im-β-Dp
	Dp-β-Py-Py-Py-Py-Py
3'-A-T-A-C-A-T-A-C-A-T-A-5'	3'-A-T-A-C-A-T-A-C-A-T-A-5'

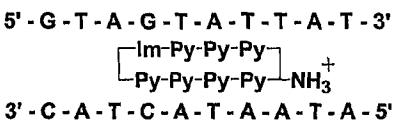
This illustrates, for example, the use of aligned sequences of monomers in the oligomeric polyamides to form antiparallel binding pairs that are capable of distinguishing all of the dsDNA base pairs from each other. The binding is described as antiparallel since the two polyamides are binding head to tail relative to each other.

As indicated above, use of folded oligomeric polyamides is preferred. The design and synthesis of folded polyamides is based on the fact that when a γ-aminobutyrate group or α-amino γ-aminobutyrate group is incorporated into the polyamide chain, this group allows the a second linear portion of the polyamide chain to fold back, form antiparallel binding pairs with a first

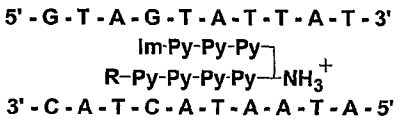
linear portion, and bind DNA in a manner similar to the 2:1 complexes. Five different folded polyamide-binding arrangements are illustrated below associated respectively with SEQ ID NO: 4, SEQ ID NO: 4, SEQ ID NO: 4, SEQ ID NO: 5, and SEQ ID NO: 6.



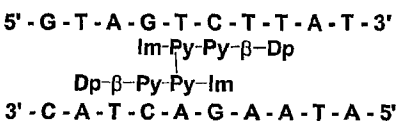
Hairpin



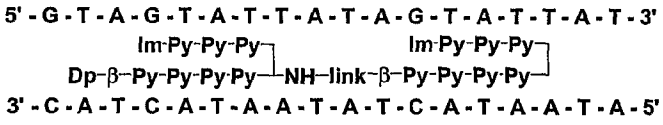
Cyclic



Amino Hairpin



H-pin



Tandem Hairpins

The folded polyamides have a number of advantages: side-by-side polyamide-DNA complexes can be formed without the entropic penalty associated with the 2:1 antiparallel polyamides; the legs (first and second oligomer portions or subchains) of the hairpin can be asymmetrical, thereby recognizing either palindromic or nonpalindromic sequences; like the 2:1 antiparallel motifs, the folded polyamide motifs have potential to differentiate all of the base pairs resulting in the potential to target more and shorter DNA sequences. Other benefits and advantages will be apparent to those skilled in the art.

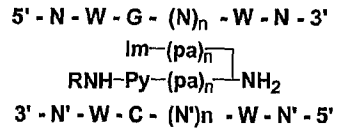
In addition, the use of tandem hairpins can be advantageous when targeting the longer range of target DNA sequences such as may be required for sufficient specificity for alleles, transgenes, and the like.

Preparation of tandem hairpins using 2,4-diaminobutyric acid derivatives as the tandem linker is described in detail in WO 98/45284, which is incorporated herein by reference for such purpose. Formation of tandem linkages generally can involve, for example, use of molecules having 2 to 12 carbon atoms, more preferably 2 to 8 carbon atoms selected from the group consisting of amino acids or carboxamides or other molecules consistent with maintaining peptide bond linkages along the polyamide backbone.

In dealing with hairpin design it will be useful to summarize certain design considerations. Where the γ-aminobutyrate group is used to form the hairpin turn, this requires a W/W (T/A or A/T) base pair at one end of the target DNA sequence. Where the N-terminus of the polyamide is an imidazole, this requires a G/C pair at the appropriate position in the target DNA sequence. Where a protonizable dimethylaminopropyl (Dp) tail is used at the C-terminus, this requires a W/W base pair at an appropriate point in the target DNA sequence. Where a β-alanine group is used to facilitate solid phase synthesis, a W/W base pair is required at an appropriate point in the target DNA sequence.

The above design aspects may result in limiting the target DNA sequence that can be targeted by hairpin polyamides to sequences having a WWG starting motif at the 5′ end of the target DNA and having a W motif at the 3′ end of the target DNA. However, if amino hairpin polyamides are used (see above), the WWG motif at the 3′ end can be shortened to a WG motif. By moving the protonatable amine from the polyamide tail to the turn in the hairpin, and correspondingly removing the amide which results from synthesis from the C-terminus tail, an important improvement is realized: the hairpin's normal

requirement for a starting sequence of WWG is shortened to just WG - see, for illustration, the hairpin associated with SEQ ID NO: 7 below.



From a DNA targeting perspective, this has the same effect as removing the b-alanine unit from the tail; however, the amino hairpins may be easier to synthesize. Moreover, as illustrated by Example VII below, the change may greatly facilitate selection of target DNA sequences.

In addition to polyamide gross structural considerations (1:1 linear, 2:1 linear, hairpin, etc.) as outlined above, molecular design for polyamide binding to target dsDNA requires consideration of selecting good dsDNA targets. To a considerable extent this can be accomplished merely by following the pairing rules shown in Table 1.

e. Pairing Rules

For linear polyamides suitable for forming 1:1 binding complexes with dsDNA, it will be sufficient simply to select a DNA sequence of interest and to design the corresponding polyamide by selecting a pyrrole for W and an imidazole for S in the dsDNA sequence. While the specificity or affinity or both of such a polyamide will not be as great as that of a hairpin or other folded structure polyamide, it is expected that such polyamides will be sufficient for sorting sperm by distinguishing X and Y chromosomes and for similar applications.

Pairing rules for polyamide structural elements useful in designing oligomeric polyamides for use with the invention are shown in the following Table.

		G/C	C/G	T/A	A/T
Im/Py	or	+	-	-	-
Im/β					
Py/Im	or	-	+	-	-
β/Im					
Hp/Py	or	-	-	+	-
Hp/β					
Py/Hp	or	-	-	-	+
β/Hp					
Py/Py		-	-	+	+
β		-	-	+	+
γ		-	-	+	+
αγ		-	-	+	+
Im		+	+	-	-
Py		-	-	+	+
Dp		-	-	+	+

TABLE 1: Pairing Rules

Referring now to Table 1, Table 1 illustrates pairing rules for designing linear polyamide 2:1 binding complexes or folded polyamide-binding complexes. Thus, a binding pair of Im/Py or Im/β is expected to bind adjacent a G/C pair in dsDNA; a binding pair of Py/Im or β/Im is expected to pair adjacent a C/G pair in dsDNA; and likewise for the other binding pairs or singlets of polyamide monomers and their complementary base pairs in dsDNA as illustrated.

Referring again to Table 1, it is indicated that Hp can be utilized to distinguish between A:T and T:A pairs in the target dsDNA. In designing polyamides for use in accordance with the invention, it is more preferred to utilize Hp in the appropriate oligomer portion for

distinguishing A:T from T:A when T closely follows (in the 3′ direction) G in the target strand since specificity and affinity have been observed to be greatest in this circumstance and the specificity of Hp for distinguishing A:T from T:A is reduced as T becomes more distant from an upstream (5′) G.

As also illustrated in Table 1, by replacing pyrrole units with β-alanine units, binding affinity for dsDNA can be retained while conferring length or conformational flexibility on the polyamide. It is preferred to utilize β-alanine units at locations where the adjacent DNA target base is A, T, or C.

In addition, as illustrated in Table 1, when designing hairpin polyamides, it is preferred to situate the hairpin at locations in the target dsDNA characterized by A/T or T/A complementary pairs. Likewise, as illustrated, it is preferred to use Dp tails at locations in the target dsDNA where A/T or T/A complementary pairs exist.

Turning again to the step of target DNA sequence selection, it will be appreciated that by using IUB-IUPAC nomenclature and the above rules and guidelines, it will be possible to specify criteria for target DNA sequence searching, including ambiguities. Then, by use of such criteria and generally available bioinformatics software, suitable target DNA sequences can be selected for use according to the different aspects of the invention.

f. Synthesis

Oligomeric polyamides for use in accordance with the invention can be synthesized by any method known in the art or hereafter developed, including without limitation synthesis methods described in the references cited in

Section 3.a. above (which are accordingly herein incorporated by reference for this purpose) or as illustrated in the Examples below, for example, by liquid phase or by solid phase methods. The preferred solid phase methods generally involve reacting a resin solid support with a resin linkage agent such as tert-butyloxycarbonylaminoacyl-pyrrole-4-(oxymethyl) phenylacetic acid, protecting and activating the pyrrole and imidazole subunits to the solid support, deprotecting the amino acids, cleaving the polyamide from the resin, and purifying the polyamide.

4. Providing Cells and Nuclei

Living and nonliving reproductive cells and nuclei of a selected species for separation in accordance with the invention may be provided using methods known in the art. For example, methodologies are well known for sperm, oocytes and ova collection, nuclei isolation, in vitro fertilization, cell suspension culture, embryo transfer and the like and these and other methods known to those skilled in the art may be used to provide a suspension of living single cells and cell clusters for contacting with oligomeric polyamides and separation in accordance with the invention.

For example, in accordance with an aspect of the invention, mammalian or avian sperm cells can be collected as is known in animal husbandry and prepared for contacting with oligomeric polyamides designed for particular target DNA sequences and binding of the polyamide to the target DNA zone of dsDNA. Semen can be collected from an avian of interest such as chickens, turkeys, ducks or geese, or from a mammal of interest such as a boar, bull, stallion, ram or any other male

mammal from which semen is collected in animal husbandry. For mammals, as an example, semen can be collected artificially using a gloved-hand method for the boar or artificial vagina for the other males listed above. Semen can also be collected from males using electro-ejaculation methods. After the semen has been collected into a collection vial, it can be diluted with an appropriate buffer that is used to extend the storage life or lifespan of the sperm outside the body. These buffers are well known and reported in the scientific literature and the chemical composition of these buffers is adapted to the species of interest. The function of a semen diluent is that of furnishing energy and nutrients to the stored sperm, provide buffering action to compensate for shifts in pH due to lactic acid formation, provide protection against rapid cooling and temperature shock, maintain the optimum osmotic pressure and balance of electrolytes for the media, inhibit the growth of microorganisms, and increase the volume of the original semen so that its use can be extended to many animals. For example, one collection of semen from a bull can be used to artificially inseminate from 300 to 800 cows and heifers. Examples of buffers for cattle include 2.9% sodium citrate - egg yolk buffer (Salisbury et al., J. Dairy Sci., 24:905 (1941)). For boar sperm, similar extenders exist as diluents for artificial insemination using fresh semen, e.g., BTS (Beltsville Thaw Solution), MR-A, Modena, and Androhep. The diluents further allow the sperm cells to be manipulated in a laboratory setting to examine sperm morphology, concentration, functionality, activity, viability, etc. Semen samples can be cryopreserved or cooled prior to contacting with specific target DNA sequence binding oligomeric polyamides following the storage period. Once the sperm

cells have been diluted to an appropriate concentration for storage or in vitro fertilization or in vivo fertilization (artificial insemination), contacting of the sperm with the oligomeric polyamide of choice can be effected by adding the chosen polyamide to a sample of the semen under conditions to effect labeling of the target dsDNA zones but which do not have any unacceptable impact on motility and/or viability.

In many instances, it will be possible to use suitably designed oligomeric polyamides which will readily permeate the cells and nuclei and bind to the target dsDNA zones. In other instances, it may be desirable to treat the cells or cellular clusters to facilitate permeation without unacceptably reducing viability of the cells. Any suitable method known to those skilled in the art may be used. These methods can include electroporation, cell-permeation-enhancing solutions, e.g., mild surfactants, and the like.

In accordance with an aspect of the invention, mammalian oocytes or eggs or embryos can be collected and prepared for contacting with target DNA-specific oligomeric polyamides. To illustrate, as is known in the animal husbandry arts, eggs or oocytes can be collected from female mammals such as cows, sows, ewes and mares for use with PA binding of DNA. Superovulation methods exist for all of the species allowing production of eggs or embryos for use in laboratory manipulation, in vitro fertilization and embryo transfer. Eggs, oocytes or embryos can also be collected from slaughtered animals. Alternatively, eggs can be produced in vitro by culture methods or embryos can be prepared by in vitro fertilization as is known in the art. Eggs or embryos can be cultured in vitro in a variety of balanced salt

solutions (i.e. TC199, M16, NCSU23 that allow for the maintenance of viability and function. The temperature of culture media used in an incubator is preferably selected to be near the body temperature of the species of interest, though other effective temperatures known in the art may also be used.

Eggs or embryos can be cultured in a wide variety of balanced salt solutions allowing one to observe developmental processes. The media can vary in composition and can be fashioned after certain manipulations used by the investigator to study mammalian developmental phenomena. For example, M2 medium is a HEPES-buffered medium useful for recovering and handling embryos out of the incubator. M16 is a CO_2-buffered medium used for culture of eggs or embryos. Both media have been modified as described in Goodall and Maro, 1986, J. Cell. Biol. 102: 568.

For culturing of eggs or embryos the medium is important but other materials and methods are also advantageous such as, for example, an incubator that can be maintained at the body temperature of the species of interest. For example, mice embryo culture is advantageously maintained at 37°C while pig embryo cultures can be maintained at 39°C for optimal development. The CO_2 monitor on the incubator should also be adjustable to reflect the optimal CO_2 environment. These and other aspects of embryo culture are well known to those skilled in the art and need not be further described here.

5. Contacting Cells and Nuclei

In accordance with the invention, following synthesis of target DNA specific oligomeric polyamide molecules as herein described, the polyamides can be

maintained in a lyophilized or other state pending use. As described above, the polyamides are also desirably conjugated with various fluorescent dyes (i.e. FITC, TRITC, BODIPY® derivatives -BODIPY® dyes are membrane soluble, for example, fluorescent fatty acid analogs available from Molecular Probes Inc. as described in, for example, US 5,338,854 and US 4,774,339 herein incorporated by reference - and the like) allowing one to visualize via fluorescence cells in which binding of oligomeric polyamides to the target DNA zone of dsDNA has occurred. Immediately prior to use, for example, the oligomeric polyamides, with or without conjugated fluorescent dyes can be resuspended in any suitable type of liquid solvent (i.e. DMSO (dimethylsulfoxide, water, etc.) resulting in a oligomeric polyamide solution. The concentration of polyamide can be varied depending upon the dilution rate. The polyamides can be stored at a variety of temperatures (-80°C to 25°C) and maintain their function of binding DNA in living cells.

In accordance with the invention, cell-permeant target DNA specific polyamide molecules can be incubated with a suspension of sperm, oocytes or ova or embryos under conditions effective for permitting the polyamide molecules to permeate the cell and nuclear membranes and to binding to the target DNA zones of dsDNA. As an example, an amount of polyamide stock solution can be added to the sperm suspension, or egg/oocyte/embryo culture droplet, to reach a final concentration of about 0.1 to 25uM. The PA-sperm or PA-egg/oocyte/embryo contact fluid can then be incubated for an effective period, for example, in the range of 30 min to 24 hour or even more to effect uptake into the cell nucleus and selective binding of the PA to target DNA.

The temperature of incubation can be at effective temperatures between about the thermotropic phase transition temperature T_m of the membranes of the sperm being sorted up through room temperature (about 23°C) to less than about 30°C or even up to less than about 39°C. The cells can be maintained in an appropriate buffer solution for maintaining cell viability, for example, for sperm, in an extender buffer such as phosphate based saline at a pH in the range of about 6.8 to about 7.6. Other media and buffers are well known to those skilled in the art and can also be used. To illustrate, an amount of oligomeric polyamide stock solution (i.e., polyamide resuspended in a suitable liquid such as water, ethanol, acetone, or the like) can be added to a cell suspension or sperm suspension or egg/oocyte/embryo suspension to reach a final concentration in the range from 0.1 to 100 μM. The polyamide solution or dry polyamides can then be added to the cell suspension and the resulting suspension can be stored at a variety of temperatures to effect uptake into the cells. For instance, storage of the contacted suspension can be maintained for a period in the range from 1 minute up to 144 hours or longer to effect uptake of the oligomeric polyamide into the cell nucleus. The exact time required will vary depending on the polyamide, the target DNA sequence, and the class of cell or cell cluster but can readily be determined by the person skilled in the art. The suspension can be mixed or unmixed during uptake.

In some or even many cases, it is believed that it will not be necessary to treat the cells to facilitate uptake of the oligomeric polyamides as described herein. However, in some cases, for example, in some instances where the polyamide or polyamide-dye conjugate length, structure or possible charge characteristics or the

particular cell type make it desirable, the cells may advantageously be treated to facilitate entry of the polyamides or polyamide-dye conjugates into the cells. For example, a chemical shock step or cell-permeation-enhancing solutions may be used to facilitate polyamides uptake, for example, using glycerol or the like. Where it is desired or advantageous to use other or more stringent techniques, such treatments can include use or liposomes or many of the techniques that are used by those skilled in the art to introduce genes or chromosomes or vectors into living cells. These methods include, but are not limited to: microinjection such as used by Gordon et al. 1980, Proc. Natl. Acad. Sci.: 7380-7384 and since extended to rabbits, sheep, cattle and pigs; electroporation; DEAE-dextran-mediated transfer; coprecipitation with calcium phosphate, and other techniques. Electroporation is a process in which cells in suspension are mixed with polyamides to be transferred. The cell-polyamide mixture is exposed to a high-voltage electric field. This treatment creates pores in the cell membrane allowing the passage of polyamide molecules into the cell interior and thence into the nucleus. The reclosing of the membrane is controlled by time and temperature and therefore by delaying this reclosing is achieved a greater probability that the polyamide molecules would enter the cell. For DEAE-dextran-mediated transfer, a DEAE dextran mixture is prepared and the polyamides are added, mixed and then transferred into the cells. The chemical events that lead to the cellular uptake are not fully understood. However, this method has been reported to yield transfection efficiencies for DNA of up to 80% (Calos et al., 1983. Proc. Nat. Acad. Sci. 80:3015.) Co-precipitation of the selected polyamides with calcium phosphate to produce

precipitates that are then added to the cells of interest can also be used. Using this method the polyamides entering the cell can be taken up into phagocytic vesicles (cf. Graham and van der Eb, 1973. Virology 52:456 - quantities of DNA taken up allow for a high frequency of genetic transformation). An option of this method is the addition of a chemical shock step to the transfection protocol using, for example, either DMSO or glycerol to facilitate uptake of the polyamides.

6. Separating Chromosomes, Nuclei, Cells and Clusters Labeled with Oligomeric Polyamides

a. Separating Chromosomes by Flow Cytometry

This is a method that provides a high-resolution analysis of, for example, the mammalian karyotype based on accurate measurement of the fluorescence of a suitably stained suspension of metaphase chromosomes. A flow cytometer is capable of measuring fluorescence with a 1% coefficient of variation and with this accuracy it is possible to differentiate between many of the chromosomes based on a quantitatively binding DNA dye. Chromosomes can be analyzed at rates up to 1000 per second and therefore a large representative number of measurements can be accumulated rapidly.

To prepare a cell suspension of metaphase chromosomes, a culture of growing cells can be treated with an agent such as colchicine or vinblastine in order to arrest sufficient cells in metaphase. The cells can be subjected to hypotonic swelling (i.e. 0.075 M KCl at 4°C for 30 minutes) and treated with a detergent to assist cell lysis. Detergents would include such as, for example, Triton X or Digitorin. The final step can be to lyse the mitotic cells so as to allow release of the mitotic chromosomes. Cell lysis can be accomplished by mechanical disruption via vortexing, passage of the cells

through a fine gauge needle or by ultrasound. After the chromosomes are isolated, they can be treated with a fluorescent nuclear dye such as ethidium bromide, Hoechst 33258, propidium iodide, chromomycin A3 and DAPI and then measured using a flow cytometer. A dual beam flow cytometer allows chromosomes to be double stained (e.g., using Hoechst 33258 and chromomycin A3) each which can be excited independently and the resultant fluorescence values can be presented as two-dimensional histograms using contour plotting. At a level of 1-2% coefficients of variation, it is possible to measure the DNA content of normal and abnormal karyotypes. Flow analysis has shown that certain chromosomes are relatively invariant in DNA content whilst others (i.e. Y chromosome) can vary considerably in DNA content in normal individuals. (Malcom, S et al. 1986, Specialist Techniques in Research and Diagnostic Clinical Cytogenetics *in* Human Cytogenetics: A Practical Approach, Rooney and Czepulkowski eds. IRL Press Ltd.)

The flow cytometer can be used for sorting X- and Y-chromosomes obtained from lysed cells. Once the chromosomes are separated, a combinatorial or composite library of polyamides can be screened for X or Y chromosome specificity. Those polyamides that are found to be X or Y specific may then be used for producing gender-enriched sperm.

b. Flow Cytometry of Nuclei, Cells and Cell Clusters

Using the methods described above, oligomeric polyamides are designed and constructed specifically for binding to target DNA sequences of interest, for example, target DNA sequences indicative of a particular allele, gene, trait, locus, or other feature of chromosomal DNA, and are then contacted with nuclei or killed or apoptotic or living cells under effective conditions for

fluorescent labeling of the target DNA sequences. The resulting labeled nuclei, cells are then separated, for example, based on the resulting differences in fluorescence to produce subpopulations of cells enriched or depleted in the target DNA sequence.

In accordance with the invention, cells characterized by association of cell-permeant target DNA specific polyamides with target DNA can be identified by detecting the fluorescence or quenching of fluorescence that indicates the association. According to one aspect of the invention, the cells or embryos can be identified and sorted using fluorescent microscopy. According to another preferred aspect of the invention, cells characterized by fluorescence resulting from oligomeric polyamide binding to target DNA molecules can be detected and sorted using flow cytometry or microfluorometry techniques.

Preferably the flow cytometry techniques are such as not adversely to affect either motility or viability of the living cells, as they are being analyzed and sorted. Such techniques are described, for example, in US 5,135,759 and US 5,985,216 that are incorporated herein by reference for this purpose.

In accordance with aspects of the invention relating to sorting sperm, ova or oocytes or embryos based on differential fluorescence resulting from selective binding of fluorescent polyamides or polyamide-dye conjugates to target DNA sequences, a flow cytometry system such as that illustrated in Figure 2 may be used.

Referring now to Figure 2 in detail, Figure 2 illustrates schematically a flow cytometry system such as may be used for effecting separations based on fluorescence in accordance with the invention. The illustrated flow cytometry system includes preparation

zone A, cytometry zone B, collection zone C, transfer zone D, and storage zone E all under automated control by controller F.

In preparation zone A as illustrated, a supply of sperm or sperm nuclei indicated at 102 and a diluent indicated at 104 are provided to constant temperature mixing zone 106 to provide diluted sperm or nuclei which can be dispensed into containers 110 on rotating table 108 for sequential positioning and delivery, for example, via line 120 to cytometry zone B.

Cytometry zone B illustrates a conventional flow cytometry system in which sample fluid provided by line 120 and sheath fluid via line 124 are introduced into nozzle 126 controlled by droplet transducer 128, for example, an ultrasonic droplet transducer, to produce droplets 152 containing predominantly only one cell or cell cluster per droplet. Laser 130 provides laser excitation 132 and 136 via filter 134 to induce differential fluorescence in cells depending on the presence or absence of fluorophores therein. Filters 146 and 148 on detector 150-focus fluorescence 144. Filter 140 on detector 142 focuses scattered light 138.

As droplets 152 leave nozzle 126, deflector plates 154 and 156 cause the droplets that have a negative charge proportional to fluorescence to be preferentially sorted into streams 174 and 176 and thence into collection vials 178 and 180 on turntable 170 of collection zone C. Vials 190 enriched or depleted in a target DNA sequence of interest are then moved via transfer zone to storage zone E where the individual samples are cryopreserved and maintained.

Referring again to FIGURE 2, reference numeral 126 illustrates the use of a flow cytometer nozzle in the methods of the invention. As those skilled in the flow

cytometry arts will appreciate, the nozzle must be sized appropriately for the class of cells or cellular cluster of interest. Such sizing is a matter of ordinary skill and need not be further described here. For separation of sperm that characteristically have flattened heads, it has been found advantageous to use nozzles that orient the sperm prior to detection. For example, a tapered needle can be used or a specially designed nozzle such as that illustrated in Rens et al., US 5,985,216 which is incorporated herein by reference, with particular reference to Figs. 1, 2 and 3 and corresponding text.

In view of advancements made in the in-vitro production of embryos from livestock and other species, it is also within the scope of the invention to use oligomeric polyamides for producing subpopulations of zygotes or embryos enriched in one or another other sex chromosome or enriched in a marker or allele of interest. For example, to effect in vitro production of embryos, immature oocytes may be removed from the ovaries, matured in vitro (IVM), and fertilized in vitro (IVF). These embryos are capable of forming live young following embryo transfer (ET) into suitable recipient females. At the time of recovery of immature oocytes from the ovary, the oocytes are arrested in meiotic prophase and this block discontinues after isolation. By identifying target DNA sequences and labeling with oligomeric polyamides as described herein, oocytes can be selected to be enriched in the chromosomal marker or allele of interest.

Oligomeric polyamides can also be used in accordance with the invention to select embryos themselves based on sex chromosomes, markers or alleles of interest. The embryos are preferably produced by IVF, then contacted with an oligomeric polyamide specific for a chromosome,

marker or allele of interest, and then the resulting embryos sorted, for example, by microscopic manipulation or flow cytometry to produce a subpopulation of embryos enriched in the chromosomal DNA feature of interest and the selected embryos then introduced by embryo transfer into suitable recipient females.

7. Iteration of Selection, Design and Testing

It will be apparent to the person skilled in the art that it may be necessary to repeat one or more of the steps illustrated in FIGURE 1 and FIGURE 1A or even FIGURE 2 to achieve optimum results. For example, in the absence of complete information about the genome of the selected class of cells, target DNA sequences albeit based on available genomic data may prove to be insufficiently specific to a chromosomal DNA characteristic of interest to permit separation, requiring an iteration of steps 20, 30, 40 and 50 of FIGURE 1 as illustrated by reference numerals 26a and 26b. Likewise, even after separation of sperm into GES, it may be further desirable to select other phenotypic characters as described herein requiring an iteration of steps 20, 30, 40, 50, and 60 as illustrated by reference numeral 26c. Similarly, referring to FIGURE 1A, steps 25 and 50′ may be iterated as indicated by reference numeral 26a′ and steps 25, 50′ and 60′ may be iterated as indicated by reference numerals 26a′ and 26c′. Other steps which may from time to time be advantageously repeated for finding or optimizing target DNA sequences or polyamides or both will be apparent to those skilled in the art from the description contained in this specification.

8. Examples

Different aspects of the invention are further illustrated and described by reference to the following Examples.

Example I: Sperm Polyamide Uptake and Viability.

This example shows that certain fluorescent polyamides and polyamide-dye conjugates permeated the cellular and nuclear membranes of sperm cells and became associated with nuclear DNA.

Boar sperm were treated with an FDA-conjugate of Im-Im-Py-Py-γ-Py-Py-Py-Py-β-Ta (target DNA sequence: wwGGwww or wwwCCww depending on orientation) as follows: 2.5 mL and 5.0 mL of sperm were combined with 5 λ of the FDA-conjugate to provide final concentrations of 6 μM and 3μM respectively. Sperm motility at outset was observed and estimated at 70%. Sperm motility was observed again after 3 days and after 6 days. Sperm motility for both treatments and controls was observed to be approximately 60% on both subsequent observations.

Aliquots (0.5ml) of pooled boar sperm (about 3.7 x 10^7 cells/mL were dispensed into plastic capped tubes and treated with varying amounts of an FITC-conjugate of Im-Im-Py-Py-γ-Py-Py-Py-Py-β-Ta to obtain final concentrations of the conjugate of 3.0, 9.0, and 18.0 μM of the conjugate. An untreated control was also employed. The samples were incubated at 18°C for up to 72 hours. The motility of all sperm groups was measured microscopically at various times. The motility of all groups at 0 and 1 hours was about 70%. After 3 days, the motility estimates for the control, 3μM, and 9μM treatment groups were about 60%, compared to about 40% for the 18μM treatment group.

Aliquots of boar sperm as described above were treated with the above described polyamide having

fluorescence labels as follows: Control, FITC, FDA, BODIPY®-FL-X, NBD-X with the results shown in the following table:

PA Molecule	Sperm Motility	Fluorescence
Control	50%	None
FITC	40%	Dead cells only
FDA	50%	Low ratio of live sperm stained
BODIPY®	50%	Higher ratio of live stained
NBD	30%	Dead cells only

These results indicated 1) sperm could be maintained for >7 days in extender buffer, 2) sperm exposed to polyamides maintained mobility similar to Control, 3) polyamide-BODIPY®-FL-X dye-conjugates showed uptake into live sperm over time.

In another study aliquots of pooled boar semen were treated as follows: Control (No polyamide-dye conjugate), polyamide-BODIPY dye conjugate #1 (Im-Py-Py-Py-β-Ta-BODIPY®-FL-X (target DNA sequence swwwww or wwGwwCww, depending on degree of overlap), polyamide-BODIPY®-FL-X dye-conjugate #2 (Im-Py-Py-γ-Im-Py-Py-Py-β-Ta-BODIPY® (target DNA sequence wwGwwCw). The results are shown in the following Table:

Treatment	0 Hours		16 Hours	
	Motility	Fluores-cence*	Motility	Fluores-cence*
Control	70%	0%	65%	0%
Conjugate #1	70%	0%	65%	90%
Conjugate #2	70%	0%	65%	25%

*relative to fluorescence of stain-saturated dead sperm

These results indicated (1) that the linear, shorter polyamide dye conjugate #1 exhibited greater uptake

(greater intensity of fluorescence) than the larger, hairpin polyamide dye conjugate #2 and bound to chromosomal DNA of live and dead sperm at about the same intensity and (2) that the larger hairpin polyamide dye molecule also permeated the cell and bound at significant levels to chromosomal DNA of living cells.

Example II: Cellular Uptake of Polyamides by Swine Sperm.

Using a polyamide (Im-Py-Py-Py-γ-Im-Py-Py-Py-β-Ta-FDA) known to have good DNA binding characteristics for its target DNA sequence and using a fluorescein-5-thiourea diacetate (FDA) label that is known to become fluorescent only after cellular uptake and hydrolysis of the acetate by cellular lipases, uptake into swine sperm was demonstrated. Motile sperm cells became fluorescent after treatment with FDA-polyamides. Confocal microscopy was used to observe the results. While all sperm cells did become fluorescent in the presence of FDA-polyamides, motility of the sperm was not affected by the polyamides. Fluorescence microscopy is also used to show that polyamides cross membranes into living sperm.

Example III: Cellular Uptake and Nuclear Localization of Polyamides by Bovine Embryos.

This example shows that fluorescent polyamide-dye conjugates can permeate the cellular and nuclear membranes of bovine embryos and become associated with nuclear DNA. Bovine eggs were fertilized in vitro and then incubated for three days. Prior to embryo culture, the embryos were treated with a polyamide-dye precursor conjugate of Im-Im-Py-Py-γ-Py-Py-Py-Py-β-Ta with FDA. Confocal microscopy was performed within a few hours of treatment. With the FDA conjugate, since the diacetate is not fluorescent, fluorescence is only observed when the FDA conjugate has entered the cells and has had the acetate groups removed by intracellular lipases producing

the FITC conjugate. By confocal microscopy, it was observed that the FDA conjugate had penetrated the embryo and the cells within, that the acetate groups were cleaved, and that the resulting FITC conjugate was concentrated in the nucleus consistently with DNA binding.

In a separate observation using a poly-lysine coated microscope slide, with 3μM as the FDA conjugate treatment, a clearly visible apparently 8-cell embryo was observed with the nuclei fluorescing due to hydrolysis of FDA to FITC and localization in the nuclei. Sperm attached to the outside of the embryo are also glowing due to FDA conjugate that has hydrolyzed.

Example IV: Cellular Uptake of Polyamides - Continued

This Example reports that various fluorescently labeled oligomeric polyamides were able to enter living motile sperm cells and localize in the nuclei. Other compounds not reported here were unable to enter the sperm under the conditions of the runs, indicating that individual polyamides and staining conditions will require testing under conditions of use including cell type, species, and polyamide dye conjugates to determine permeation and staining of nuclei of living cells.

Table: Individual Polyamide-dye conjugates taken up by intact, live sperm:

Run	Compound	Comment
1	It3(FDA)-Py-Py-β-Da	
2	It3(BOFLX)-Py-Py-β-Py-Py-Py-β-Da	
3	Im-Py-Py-Py-β-Ta-BOFLX	
4	It3(BOFLX)-Py-Py-Py-β-Ta-Cl_2	
5	Im-Py-Py-$\beta_{N(BOFLX)}$-Py-Py-Py-β-Da	
6	Im-Py-Py-Py-β-Ta-FDA	
7	Im-Py-β-Py-Py-β-Ta-	species specific - dye uptake

	BOFLX	into boar only, not bull

Example V: Cellular Uptake and Nuclear Localization of Polyamides During Cell Cycle

This Example evaluated the ability of polyamide dye conjugates to bind DNA quantitatively to somatic cell DNA.

The cell cycle is evaluated using flow cytometry and a DNA stain (such as propidium iodide) that allows quantification of cellular DNA. The cell cycle comprises interphase (DNA duplication), prophase, metaphase (chromosome alignment on the spindle), anaphase and telophase (DNA division complete). Since polyamide dye conjugates were found to bind DNA of sperm using confocal microscopy, we investigated their ability to bind DNA quantitatively as no information existed in the literature. A proliferating mammalian cell line was used to measure the polyamide dye conjugates ability to bind quantitatively throughout the cell cycle. The cells were permeabilized using 70% ethanol exposure at -20°C. The cells were pelleted and re-suspended in PBS at a concentration of 150,000-200,000 cells/ml. Ten microliters of 1mM in DMSO stock of a PA dye conjugate:

Im-Py-Py-β-Py-Py-Py-β-Ta-BOFLX

was added to the permeabilized cells and then processed using a FACS Calibur (Becton Dickinson, Mountain View, CA) flow cytometer. A control sample was also set up using propidium iodide staining of cells from the same sample. Processing cells for propidium iodide was performed as above for the polyamide dye conjugate. About 17,000 cells were processed for each sample to obtain statistical analysis for measuring DNA quantity. Single

parameter frequency histograms were generated and the results are shown below indicating that no difference existed between propidium iodide and the polyamide dye conjugate. The data therefore indicate that the polyamide dye conjugate, like propidium iodide, was capable of quantitating DNA of somatic cells at different stages of the cell cycle. This information justified moving toward evaluating the ability of polyamide dye conjugates to quantitating sperm nuclei in which all DNA exists in tightly packed nucleosomal form.

Stain	CV	Percentage of cells within cell cycle stages		
		G0/G1	G2M	Synthesis phase
Propidium iodide	4.64	56.8	15.8	27.2
GES 17-BODIPI	5.58	51.4	17.7	30.8

Using confocal microscopy, the ability of a polyamide dye conjugate to bind non-specifically to sperm DNA was evaluated using confocal microscopy. A polyamide-FITC dye conjugate (1mM in DMSO) was used to treat boar sperm (70% motility) suspended in BTS (Beltsville Thaw Solution) extender (2uM final concentration of PA-FITC). Sperm were placed on a poly-lysine coated microscope slide and covered using a cover slip. The cells were observed (60x magnification) to have a uniform fluorescence throughout the nucleus as evidenced by multiple sagittal sections through the sperm head. This indicated that the polyamide dye conjugate is capable of binding to sperm DNA non-specifically and uniformly within the sperm head.

Example VI: Target DNA Selection - FRET System

The sequence (SEQ ID NO: 8) shown below was selected from the *Bos taurus* TSPY gene which is from the Y chromosome:

```
          WWGCGGWCCW      WWGGCWGGCW
5'  ATGCGGTCCTGGGCAATTGGCAGGCT 3'     SEQ ID NO: 8
3'  TACGCCAGGACCCGTTAACCGTCCGA 5'
```

The following polyamide has been designed to target the 5'-("left") site:

Im-β-Im-Im-β-Py-Py-β-Im-Im-Py-Py-β-Im-Py-β-Ta-BODIPY-FL-X

The following polyamide has been designed to target the 3'-("right") site:

Im-Im-Py-β-Im-Im-Py-γ-Im-Py-Py-β-Im-Py-Py-β-Ta-BODIPY-TMR-X

These polyamide-dye conjugates place the FRET dyes approximately 16 bp apart.

Example VII: Target Selection

A 4.2 kb sequence of DNA (BRY.4a) is repeated many times exclusively on the bovine Y chromosome. Finding polyamide binding sites within this repeat sequence is desirable since the polyamide binding and any resulting fluorescent signal would occur many times when the Y chromosome is present. A detailed analysis of the repeat sequence looking for good polyamide binding sites has revealed the presence of the site shown below in SEQ ID NO: 9.

5' TGGAAAAGGTAAGGAAAAGTTCT 3' SEQ ID NO: 9

One method to design polyamides to bind to longer DNA sequences is by linking short hairpin polyamide modules together. Using this approach, the polyamide shown below was designed to bind to the sequence above:

M1-β-M2-β-M3-β-M4-BODIPY-FL-X

where M1 is Im-Im-Py-Py-γN-Py-Py-Py-Py-Dp

M2 is Im-Im-Py-γN-Py-Py-Py-

M3 is Im-Im-Py-Py-γN-Py-Py-Py-Py-

and M4 is Im-Py-Py-Py-γN-Im-Py-Py-Py-

Example VIII: Preparation of FITC Labeled Polyamide

12 µl of diisopropylethylamine was added to a solution of 20.07 mg of the tetra-trifluoroacetic acid salt of Im-Py-Py-Py-γ-Im-Py-Py-Py-β-$NHCH_2CH_2CH_2N(Me)CH_2CH_2CH_2NH_2$ in 2 mL of anhydrous DMF (dimethylformamide). An anhydrous DMF solution (2 ml) of fluorescein-5-isothiocyanate (6.14mg) was added and the mixture stirred overnight at room temperature. The product was isolated via reverse phase chromatography using a methanol/water gradient. Lyophilization from a t-butanol/water mixture gave 17.2 mg of Im-Py-Py-Py-γ-Im-Py-Py-Py-β-$NHCH_2CH_2CH_2N(Me)CH_2CH_2CH_2NHC(=S)NH$-5-fluorescein (74%) yield. Mass spec: $M+H^+$ (m/z=1655) and $M+2H^+$ (m/z=828) observed.

Example IX: Preparation of FDA and TMR Labeled Polyamides

In a manner similar to the method described above, the following polyamide-dye conjugates were prepared: Im-Im-Py-Py-g-Py-Py-Py-Py-$NHCH_2CH_2CH_2N(Me)CH_2CH_2CH_2NHC(=S)NH$-5-fluorescein diacetate(FDA): (78%) yield. Mass spec: $M+H^+$ (m/z=1738) and $M+2H^+$ (m/z=870) observed; Im-Py-Py-Py-g-Im-Py-Py-Py-b-$NHCH_2CH_2CH_2N(Me)CH_2CH_2CH_2NHC(=S)NH$-5-tetramethylrhodamine(TMR): (61%) yield. Mass spec: $M+H^+$ (m/z=1710) and $M+2H^+$ (m/z=855.5) observed.

Example X: Preparation of BODIPY-FL-X Labeled Polyamide

6.4 µl of diisopropylethylamine was added to a solution of 11.0 mg of the tetra-trifluoroacetic acid salt of Im-Py-Py-Py-γ-Im-Py-Py-Py-β-$NHCH_2CH_2CH_2N(Me)CH_2CH_2CH_2NH_2$ in 1 mL of anhydrous DMF (dimethylformamide). An anhydrous DMF solution (1 ml) of 6-((4,4-difluoro-5,7-dimethyl-4-bora-3a,4a-diaza-s-indacene-3-propionyl)amino)hexanoic acid, succimidyl

ester (BODIPY® FL-X, SE) (3.76 mg) was added and the mixture stirred 36 hours at room temperature. The product was isolated via reverse phase chromatography using a methanol/water gradient. Lyophilization from a t-butanol/water mixture gave 7.76 mg of Im-Py-Py-Py-γ-Im-Py-Py-Py-β-$NHCH_2CH_2CH_2N(Me)CH_2CH_2CH_2NH$-BODIPY-FL-X (61% yield). Mass spec: $M+2H^+$ (m/z=827) observed.

Example XI: Preferred Syntheses of Intermediates and of Im-Py-Py-Py-γ-Im-Py-Py-Py-β-Ta

For convenience, literature synthetic methods referenced in this Example are cited at the end of the Example. Monomer building blocks were prepared by the procedures reported in the literature. Several modifications were made during the syntheses of these building blocks and those would be listed wherever necessary.

XI.1 Synthesis of 4-[(*tert*-Butoxycarbonyl)amino-1-methylpyrrole-2-carboxylic acid. Boc protected pyrrole monomer was prepared by the literature procedures[1-4]

XI.2 2-(Trichloroacetyl)-1-methylpyrrole

To a mechanically stirred solution of trichloroacetyl chloride (820 g, 4.51 mol) in 1200 mL of anhydrous diethyl ether, was added dropwise over a period of 4 h a solution of *N*-methylpyrrole (365 g, 4.5 mol) in 1200 mL of anhydrous diethyl ether. After stirring the reaction mixture overnight at room temperature, the reaction was quenched by the dropwise addition of a solution of 328 g of potassium carbonate in 1200 mL of water in 2 h. The layers were separated, and the ether layer was dried with anhydrous sodium sulfate and concentrated *in vacuo* to afford 2-(trichloroacetyl) pyrrole as a yellow tan crystalline solid (960 g, 91% yield). The crude product was analyzed by TLC and ^1H NMR

which were interpreted to indicate that it was a pure product capable of use without further purification.

XI.3 4-Nitro-2-(trichloroacetyl)-1-methylpyrrole

To a cooled (-42 °C) solution of 2-(trichloroacetyl)-1-methylpyrrole (330.22 g, 1.458 mol) in acetic anhydride (1650 mL) in a 5 L flask equipped with a mechanical stirrer was added 125 mL of fuming nitric acid (2.81 mol) over a period of 1 h while a temperature of the reaction mixture was maintained between -42 to -44 °C. The reaction mixture was first stirred at -40 °C for 1.5 h and then stirred at -30 °C for 50 min. The reaction mixture was carefully allowed to warm to 0 °C and stirred for an additional 2h. The mixture was cooled to -30 °C and isopropyl alcohol (1650 mL) was added. The solution was stirred at -20 °C for 30 min during which time a white precipitate formed. The solution was allowed to stand for 15 min and the resulting white precipitate was filtered and dried *in vacuo* to afford 4-nitro-2-(trichloroacetyl)-1-methylpyrrole as a white crystalline solid (241.31 g, 61% yield). The product was characterized using TLC, ^{1}H NMR and mass spectroscopy and by comparison of its spectral and chromatographic data with those given in the literature[1,3].

XI.4 Methyl-4-nitro-1-methylpyrrole-2-carboxylate

To a solution of 4-nitro-2-(trichloroacetyl)-1-methylpyrrole (805 g, 2.965 mol) in 2000 mL of methanol in a 4 L Erlenmeyer flask equipped with a mechanical stirrer was added dropwise a 25 wt. % solution of sodium methoxide in methanol (5.7 mL, 0.0249 mol). The reaction mixture that turned yellow with an exothermic reaction was stirred at room temperature. The reaction mixture started to thicken with the solid formed. The reaction mixture was quenched by the addition of concentrated

sulfuric acid (2.5 mL). The reaction mixture was diluted with additional methanol (3000 mL) and then heated to reflux until all solid dissolved. The reaction mixture was allowed to slowly cool to room temperature as methyl-4-nitro-1-methylpyrrole-2-carboxylate crystallized as colorless needles, which were filtered and dried *in vacuo* (467.4 g,). The mother liquor was concentrated, diluted with water and was extracted with methylene chloride. The combined methylene chloride extracts were extracted with 5% aqueous sodium bicarbonate solution and dried with anhydrous $MgSO_4$. The solvent was removed *in vacuo* and the residue was recrystallized from methanol to afford additional 55.23 g of the product (522.63 g, 95.7% total yield). The product was characterized using TLC, ^1H NMR and mass spectroscopy and by comparison of its spectral and chromatographic data with those reported in the literature[1,3].

XI.5 Methyl-4-amino-1-methylpyrrole-2-carboxylate

The reduction of the methyl-4-nitro-1-methylpyrrole-2-carboxylate was performed by a transfer hydrogenation[6-8] with ammonium formate and Pd/C. To a solution of methyl 4-nitro-1-methylpyrrole-2-carboxylate (80 g, 0.435 mol) in ethyl acetate (1500 mL) was added ammonium formate (218 g, 3.53 mol) and 10 % Pd/C (75 g) and the reaction mixture was refluxed under a nitrogen atmosphere for 45 min. The reaction mixture was filtered through a bed of Celite, washed with ethyl acetate (2 X 50 mL). The solvent was evaporated under reduced pressure to afford methyl-4-amino-1-methylpyrrole-2-carboxylate as a dirty yellow oil (66 g, 98 %). The crude product was analyzed by TLC and ^1H NMR which were interpreted to indicate that it was a fairly pure product capable of use without further purification.

XI.6 Methyl-4-[(*tert*-butoxycarbonyl)amino]-1-methylpyrrole-2-carboxylate and 4-[(*tert*-Butoxycarbonyl)amino]-1-methylpyrrole-2-carboxylic acid

Methyl-4-amino-1-methylpyrrole-2-carboxylate was prepared by the reduction of methyl-4-nitro-1-methylpyrrole-2-carboxylate (80.17 g, 435.3 mmol) by transfer hydrogenation with ammonium formate (164.40 g, 2607 mmol) and 10% Pd on Carbon (50% water wet) (11.57g, 5.4 mmol) in 1.5 L ethyl acetate. The solvent was removed *in vacuo* to give methyl-4-amino-1-methylpyrrole-2-carboxylate as an almost colorless oil. The oil was dissolved in 50 mL of anhydrous THF. To a mechanically stirred solution of the oil was added dropwise over a period of 10 min a solution of di-*tert*-butyldicarbonate (120 mL, 522 mmol) in anhydrous THF. The reaction mixture was stirred at room temperature for 1 h. The solvent was removed *in vacuo* to afford crude methyl 4-[(*tert*-butoxycarbonyl) amino]-1-methylpyrrole-2-carboxylate as a colorless solid contaminated with di-*tert*-butyldicarbonate as indicated by ^{1}H NMR.

To a mechanically stirred suspension of the solid from above in absolute ethanol (1000 mL) was added a 2.5 M sodium hydroxide solution in water (500 mL, 1250 mmol)and the reaction mixture was refluxed for 2 h. After completion of the reaction, ethanol was removed *in vacuo* and the residue was dissolved in water (200 mL), then extracted with ethyl acetate. The aqueous layer was acidified (pH ~ 1-2) with a cooled solution of conc. hydrochloric acid (112 mL) in water (250 mL) by dropwise addition and maintaining the reaction temperature less than 15 °C. The reaction mixture was extracted several times with ethyl acetate and the combined extract was dried with anhydrous $MgSO_4$, filtered and stripped to

afford 4-[(*tert*-butoxycarbonyl)amino]-1-methylpyrrole-2-carboxylic acid as a colorless solid (108.87 g, 98% overall yield). The product was characterized using TLC, ^{1}H NMR and mass spectroscopy and by comparison of its spectral and chromatographic data with those reported in the literature[1,3].

XI.7. Synthesis of 4-[(*tert*-Butoxycarbonyl)amino-1-methyl-imidazole-2-carboxylic acid

Boc protected imidazole monomer was prepared by the literature procedures.[1,3-5]

XI.8. 2-(Trichloroacetyl)-1-methylimidazole

To a mechanically stirred solution of trichloroacetyl chloride (555 g, 3.05 mol) in 1800 mL of methylene chloride was added dropwise over a period of 5 h a solution of *N*-methylimidazole (250 g, 3.05 mol) in 750 mL of methylene chloride at room temperature. After stirring the reaction mixture overnight at room temperature, the reaction mixture was cooled to 0 °C (ice/salt bath) and triethylamine (425 mL, 3.05 mol) was added dropwise over a period of 2 h and stirred overnight at room temperature. The reaction mixture was extracted with water and methylene chloride solution was dried with anhydrous $MgSO_4$. The solvent was removed *in vacuo* to afford crude 2-(trichloroacetyl)imidazole as a yellow solid. Recrystallization of the residue was accomplished by extracting the residue several times with boiling hexane (1500 mL) and cooling the hexane extracts to room temperature afforded a pale yellow crystalline solid (558.5 g, 81% yield). The product was characterized using TLC, ^{1}H NMR and mass spectroscopy and by comparison of its spectral and chromatographic data with those reported in the literature.[1,4]

XI.9. 4-Nitro-2-(trichloroacetyl)-1-methylimidazole

Fuming nitric acid (60 mL, 1.3 mol) was added dropwise to acetic anhydride (730 mL) at 0 °C and subsequently conc. Sulfuric acid (1.5 mL) was added to the reaction mixture with stirring. Then powdered 2-(trichloroacetyl)-1-methylimidazole (100 g, 440 mmol) was added gradually to the stirred solution over a period of 1 h while a temperature of the reaction mixture was maintained at 0 °C. The reaction mixture was carefully allowed to warm to room temperature and stirred overnight at room temperature. The reaction mixture was poured onto ice and extracted with chloroform. The chloroform extract was concentrated *in vacuo* to remove excess acetic anhydride and acetic acid. The resulting solid was slurried in a small volume of chloroform to ensure that the by product 5-nitro-2-(trichloroacetyl)-1-methylimidazole was dissolved. Trituration with an 1:1 ether/hexane mixture gave very pure 4-nitro-2-(trichloroacetyl)-1-methylimidazole as a colorless crystalline solid (75.4 g, 63% yield). The product was characterized using TLC, ^{1}H NMR and mass spectroscopy and by comparison of its spectral and chromatographic data with those reported in the literature[1,4].

XI.10. Methyl-4-nitro-1-methylimidazole-2-carboxylate

To a solution of 4-nitro-2-(trichloroacetyl)-1-methylimidazole (230 g, 844 mmol) in 767 mL of methanol was added dropwise a 0.5 M solution of sodium methoxide in methanol (17 mL, 8.5 mmol) at room temperature. The reaction mixture was stirred at room temperature, and concentrated *in vacuo* to remove most of the methanol. The residue was dissolved in chloroform (1000 mL), washed with water, and the chloroform layer was dried with anhydrous $MgSO_4$. Most of the chloroform was removed *in vacuo* and the mixture was diluted with ether to afford methyl-4-nitro-1-methylimidazole-2-carboxylate (149.87 g,

96% yield). The product was characterized using TLC, ^{1}H NMR and mass spectroscopy and by comparison of its spectral and chromatographic data with those reported in the literature.[1]

XI.11. Methyl-4-amino-1-methylimidazole-2-carboxylate

The literature procedure[1] for the reduction of methyl-4-nitro-1-methylimidazole-2-carboxylate by the hydrogenation with hydrogen and Pd/C also did not work very well in spite of following the literature procedures and conditions. The reduction of the methyl-4-nitro-1-methylimidazole-2-carboxylate was performed by the following superior procedure, a transfer hydrogenation[5-8] with ammonium formate and Pd/C (as described above for the reduction of the methyl-4-nitro-1-methylpyrrole-2-carboxylate).

To a solution of methyl 4-nitro-1-methylimidazole-2-carboxylate (10 g, 54 mmol) in ethyl acetate (100 mL) was added ammonium formate (17 g, 270 mmol) and 20 % Pd hydroxide on carbon (Pearlman's catalyst) (0.5 g) and the reaction mixture was refluxed under a nitrogen atmosphere for 45 min. The reaction mixture was filtered through a bed of Celite, washed with ethyl acetate (2 X 50 mL). The solvent was evaporated under reduced pressure to afford the crude product as a dirty yellow solid (7.629 g, 91%). The crude product was analyzed by TLC and ^{1}H NMR which were interpreted to indicate that it is a fairly pure product to be used without further purification.

XI.12. Methyl 4-[(*tert*-butoxycarbonyl)amino]-1-methylimidazole-2-carboxylate

To a stirred solution of methyl 4-[(*tert*-butoxycarbonyl)amino]-1-methylimidazole-2-carboxylate (23 g, 148.4 mmol) in a mixture of water (46 mL) and acetone (115 mL) was added di-*tert*-butyl dicarbonate (48.6 g, 222.7 mmol). A homogeneous solution was obtained after

stirring the reaction mixture overnight at room temperature. The reaction mixture was concentrated *in vacuo* to remove most of the acetone, then the residue was dissolved in $CHCl_3$, washed with water and the $CHCl_3$ layer was dried with anhydrous $MgSO_4$. Evaporation of the solvent *in vacuo* afforded an oil which on trituration with hexane afforded an off white solid (34.65 g, 92% yield). The product was characterized using TLC, ^1H NMR and mass spectroscopy and by comparison of its spectral and chromatographic data with those reported in the literature[1].

XI.13. 4-[(*tert*-Butoxycarbonyl)amino]-1-methylimidazol-2-carboxylic acid

To a magnetically stirred solution of lithium hydroxide monohydrate (10.37 g, 247.14 mmol) in 95% methanol (350 mL) was added methyl 4-[(*tert*-butoxycarbonyl)amino]-1-methylimidazole-2-carboxylate (51.5 g, 202 mmol). The yellow orange solution was stirred overnight at room temperature. Chloroform (350 mL) was added to above mixture, neutralized with 1N hydrochloric acid (247 mL, 248.35 mmol) and after stirring for 15 min chloroform layer was removed. The chloroform layer was washed with water and dried with anhydrous sodium sulfate and filtered. Evaporation of the solvent *in vacuo* afforded the 4-[(*tert*-butoxycarbonyl)amino]-1-methylimidazole-2-carboxylic acid product as a pale yellow solid (47.5 g, 97.5% yield). The product was characterized using TLC, ^1H NMR and mass spectroscopy and by comparison of its spectral and chromatographic data with those reported in the literature[1].

XI.14. Syntheses of Boc-protected Dimers

Two Dimeric building blocks: 4-[[[3-[(*tert*-Butoxycarbonyl)amino]propyl]carbonyl]-amino-1-

methylimidazole-2-carboxylic acid and 4-[[[4-[(*tert*-Butoxycarbonyl) amino]-1-methylpyrrol-2-yl]carbonyl]amino-1-methylimidazole-2-carboxylic acid were prepared by the procedures reported in the literature.[1] The products were characterized using TLC, ^1H NMR and mass spectroscopy and by comparison of their spectral and chromatographic data with those reported in the literature[1].

XI.15. Solid Phase Syntheses of Polyamides

Solid phase synthesis of polyamide was performed on an ABI 433A peptide synthesizer on a 0.26 mmol scale using 1 gm of BOC-β-ala-Pam Resin (0.26 mmol/g) and four equivalents of monomer building block (1.0 mmol). Each cycle of amino acid addition involved deprotection of BOC-β-ala-Pam Resin (1.0 g, .26 mmol, placed in a 41 mL reaction vessel) with 25% TFA/CH_2Cl_2 for 3 min, draining the reaction vessel, and then deprotection with 50% TFA/CH_2Cl_2 for 16 min, and washing with DCM (4 X 7 mL). The amino acid monomer (1.0 mmol) was weighed into a synthesis cartridge, dissolved in 3 mL DMF and 2 mL DIEA, allowed to mix in the cartridge for 3 min and then the dissolved amino acid was transferred to the activator vessel. The activation of the amino acid was performed with 1 mmol HBTU in DMF (2 mL), transferred directly to the activator vessel and was allowed to stand for 10 min. The deprotected β-ala-Pam Resin was neutralized with 10% solution of DIEA in DCM and neutralized again with 10% DIEA in DMF and then washed with DMF (6 X 5 mL). After deprotection and neutralization of the resin, the activated amino acid monomer solution from activator was then transferred to the reaction vessel and allowed to couple for 30 min, and a solution of DMSO/NMP was added and coupling was then continued for another 45 min (total coupling time 75 min). DIEA (1 mL) was added and allowed

to couple for another 60 min, the reaction vessel was drained and the resin was washed with DMF, capped with 10% acetic anhydride, 5% DIEA in DMF for 9 min, the resin was washed again with DCM, and the reaction vessel was drained and the process was repeated for each amino acid monomer.

The ABI 433 peptide synthesizer was left in the standard hardware configuration for BOC chemistry protocols. Reagent bottle positions 1 and 7 were DIEA, reagent bottle position 2 was TFA, reagent bottle position 4 was acetic anhydride, reagent bottle position 5 was DMSO/NMP, reagent bottle position 6 was methanol (for washing only), reagent bottle position 8 was HBTU/DMF, reagent bottle position 9 was DCM and reagent bottle position 10 was DMF.

XI.16. Cleavage of the resin

1 gm of the resin was treated with 6 mL of 3-(dimethylamino)propylamine (DA) or 3,3'-diamino-*N*-methyldipropylamine (TA) at 45 °C for 18 h. The resin was removed by filtration and washed with 12 mL of water. The filtrate was analyzed by analytical HPLC at 254 and 304 nm.

XI.17. Purification of Polyamides

The filtrate was evaporated on the rotary evaporator to remove water and Dp or Ta. The residue was dissolved in a 1:1 mixture of DMF and water (10-15 mL), filtered through a 0.45 μm nylon filter. The filtrate was loaded on a C_{18} preparative HPLC column and the polyamide was eluted with a linear gradient of 20-80% methanol in water containing 0.1% TFA. The fractions collected were analyzed by analytical HPLC to verify their purity. The appropriate fractions were mixed together and evaporated. The residue was dissolved in a mixture of *tert*-butanol and water (25-50%) and lyophilized to afford the

polyamide as a cream to yellow powder. Preparative HPLC purification afforded the polyamide with purity greater than 98%.

XI.18. Characterization of Polyamides

The polyamides were characterized by low resolution liquid chromatography-mass spectrometry (LC-MS) and high-resolution mass spectrometry (HRMS).

XI.19. Synthesis of Im-Py-Py-Py-γ-Im-Py-Py-Py-β-Dp

Im-Py-Py-Py-γ-Im-Py-Py-Py-β-Pam-resin was prepared in a stepwise manner on a peptide synthesizer by automated solid phase synthesis protocols described above. A sample of Im-Py-Py-Py-γ-Im-Py-Py-Py-β-Pam-resin (0.307 g, 0.26 mmol/g) was placed in a 25 mL glass vial, treated with 3-(dimethylamino)propylamine (6 mL), and heated at 45 °C with stirring for 18 h. The resin was removed by filtration and washed with 12 mL of water. The filtrate of the crude sample was analyzed by analytical HPLC at 254 and 304 nm. The filtrate was evaporated *in vacuo* and the residue was dissolved in a 1:1 mixture of DMF and water (10-15 mL), filtered through a 0.45 μm nylon filter. The filtrate was loaded on a C_{18} preparative HPLC column and the polyamide was eluted with a linear gradient of 20-80% methanol in water containing 0.1% TFA. The fractions collected were analyzed by analytical HPLC to verify their purity. The appropriate fractions were mixed together and evaporated. The residue was dissolved in a mixture of *tert*-butanol and water (1:2) and lyophilized to afford the polyamide as a cream powder (28.4 mg). The polyamide was characterized by ^{1}H NMR and mass spectroscopy.

XI.20. References for Synthesis Example XI

The following references, which are incorporated herein by reference, were cited in this Example.

1. E.L. Baird and P.B. Dervan, *J. Am. Chem. Soc.*, *118*, 6141-6146 (**1996**).
2. D. M. Bailey, R.E. Johnson and N.F. Albertson, *Org. Synth.*, *51*, 100-102 (**1971**).
3. E. L. Baird and P.B. Dervan, PCT **US 97/003332 ; WO 98/49142**
4. E. Nishiwaki, S. Tanaka, H. Lee, and M. Shibuya, *Heterocycles*, *27*, 1945-1952 (**1988**).
5. L. Grehn, L. Ding, and U. Ragnarsson, *Acta. Chem. Scand. 44*, 67 (1990).
6. S.Ram and R.E. Ehrenkaufer, *Synthesis*, 91-95 (**1988**).
7. S. Ram and R.E. Ehrenkaufer, *Tetrahedron Lett.*, *25*, 3415-3418 (**1984**).
8. B. König and M. Rödel, *J. Chem. Soc., Chem. Comm.*, 605-606 (**1998**)

Example XII: Synthesis of Hydroxypyrrole and Intermediates

XII.1 Ethyl Sarcosinate

The liberation of the free ester from its hydrochloride salt was accomplished by passing dry ammonia gas through a suspension of sarcosine ethyl ester hydrochloride (50 g, 325.5 mmol) in diethyl ether (600 mL) at 0 °C (ice-bath) for 3 h. Precipitated ammonium chloride was removed by filtration and washed with ether. The filtrate was concentrated, first by rotary evaporation and then on a vacuum pump for 30 min afforded ethyl sarcosinate as a pale pink liquid (38.55 g, 100 %).

XII.2: Ethyl ethoxymethylenenitroacetate

A mixture of ethyl nitroacetate (30 g, 225.4 mmol), triethyl orthoformate (75 mL, 450 mmol) and acetic anhydride (61 mL, 640 mmol) were heated for 1 h at 120 °C and for a further 1 h at 130 °C in a distillation apparatus. The reaction mixture was finally heated at 140 °C and the distillate (~ 100 mL) was collected. The excess of triethyl orthoformate and acetic anhydride was removed under reduced pressure to give a dark orange

viscous liquid (41.4 g). Fractional distillation of the crude product under reduced pressure (105-106 °C /0.2 mm Hg) afforded a light yellow liquid (33.87 g, 79 %). GC and ^{1}H NMR analysis of the purified product indicated the presence of Z,E-isomers in a 73:27 ratio.

XII.3 Ethyl-N-[(2-(nitro-2-ethoxycarbonyl)vinyl] sarcosinate

Ethyl sarcosinate (3.30 g, 28.17 mmol) was added to ethyl ethoxymethylenenitroacetate (5.26 g, 27.80 mmol) at room temperature, an exothermic reaction occurred and a dark orange red solution was obtained. The reaction mixture was stirred at room temperature for 2 h. Ethanol was removed by rotary evaporation to afford an orange red oily liquid (7.186 g, 99 %): MS (EI, 70 eV) m/z (relative intensity) 261 (M+1, 1), 260 (M^+, 12), 214 (10), 187 (37), 159 (100), 142 (24), 113 (41), 85 (72), 69 (17), 42(70).

XII.4: Ethyl 1-methyl-4-nitro-3-hydroxypyrrole-2-carboxylate

To a solution of ethyl-N-[(2-nitro-2-ethoxycarbonyl)-vinyl]sarcosinate (32.12 g, 123.42 mmol) in abs. ethanol (150 mL) was added a solution of sodium ethoxide prepared from sodium (4.0 g, 174.0 mmol) and abs. ethanol (100 mL). The reaction mixture was refluxed under a nitrogen atmosphere for 2 h to afford a dark orange-brown solution. Ethanol was removed under reduced pressure to give a dark brown residue. The residue was dissolved in water (500 mL) and acidified with 20 % H_2SO_4 to give a brown precipitate, filtered and washed with water. The filtrate was extracted with CH_2Cl_2 (10 X 100 mL). The brown solid was dissolved in CH_2Cl_2 and the solution was mixed with CH_2Cl_2 extracts. The combined extracts were dried with anhydrous Na_2SO_4. Filtration and rotary evaporation gave an orange brown solid (11.28 g). The residue was dissolved in CH_2Cl_2 (250 mL) and passed through a pad of Silica gel (100 g), washed several times with CH_2Cl_2, evaporation of the solvent gave an orange crystalline solid (5.60 g, 21.2 %). Recrystallization from ethyl acetate/hexane afforded very pale yellow

needles; mp 126-128 °C; HRMS: Calcd for $C_8H_{11}N_2O_5$ [M+H$^+$]: 215.0668, Found; 215.0666. Anal. Calcd for $C_8H_{10}N_2O_5$: C, 44.86; H, 4.70; N, 13.07. Found: C, 44.97; H, 4.76; N, 12.97.

XII.5: Ethyl 1-methyl-4-amino-3-hydroxypyrrole-2-carboxylate

To a solution of ethyl 1-methyl-4-nitro-3-hydroxypyrrole-2-carboxylate (0.90 g, 4.20 mmol) in ethyl acetate (50 mL) was added ammonium formate (2.64 g, 41.86 mmol) and 10 % Pd/C (1.20 g) and the reaction mixture was refluxed under a nitrogen atmosphere for 30 min. The reaction mixture was filtered through a bed of Celite, washed with ethyl acetate (2 X 50 mL). The solvent was evaporated under reduced pressure to afford the crude product as a dirty yellow solid (0.664 g, 86 %); MS (EI, 70 eV) m/z (relative intensity) 185 (M+1, 3), 184 (M$^+$, 35), 139 (16), 138 (100), 137 (16), 110 (25), 109 (9), 82 (9), 55 (9), 42 (10).

XII.6 Ethyl 1-methyl-4-nitro-3-methoxypyrrole-2-carboxylate

A mixture of ethyl 1-methyl-4-nitro-3-hydroxypyrrole-2-carboxylate (0.071 g, 0.3315 mmol), dimethyl sulfate (0.055 g, 0.436 mmol) and anhydrous potassium carbonate (0.5 g, 3.90 mmol) in dry acetone (20 mL) was refluxed under a nitrogen atmosphere for 22 h. Potassium carbonate was removed by filtration and the filtrate was evaporated under reduced pressure to afford an orange crystalline solid. Excess dimethyl sulfate was destroyed by treating the residue with aqueous ammonia (1 mL). Water (2 mL) was added and the residue was extracted with ether (30 mL). The combined ether extracts were washed with aqueous sodium hydroxide solution and water. Removal of ether under reduced pressure afforded a cream residue. The residue was dissolved in CH_2Cl_2 (25 mL), dried with anhydrous Na_2SO_4, filtered and evaporated under reduced pressure to yield a cream microcrystalline solid (0.067 g, 89 %). Recrystallization from acetone/hexane afforded colorless needles, mp 145-147 °C; HRMS (M+H):

Calcd for $C_9H_{13}N_2O_5$ [M+H$^+$]: 229.0824, Found 229.0834. Anal. Calcd for $C_9H_{12}N_2O_5$: C, 47.37; H, 5.29; N, 12.27. Found. C, 47.44; H, 5.36; N, 12.28.

XII.7: Ethyl 1-methyl-4-amino-3-methoxypyrrole-2-carboxylate

To a solution of ethyl 1-methyl-4-nitro-3-methoxypyrrole-2-carboxylate (3.34 g, 14.64 mmol) in ethyl acetate (150 mL) was added ammonium formate (9.23 g, 146.37 mmol) and 10 % Pd/C (4.5 g) and the reaction mixture was refluxed under a nitrogen atmosphere for 30 min. The reaction mixture was filtered through a bed of Celite, washed with ethyl acetate (50 mL). The solvent was evaporated under reduced pressure to afford the product as an orange viscous liquid (2.753 g, 95 %); MS (EI, 70 eV) m/z (relative intensity) 199 (M+1, 5), 198 (M$^+$, 41), 183 (8), 137 (100), 109 (6), 81 (5), 54 (6), 42 (8).

XII.8: Ethyl 1-methyl-4-(tert-butoxycarbonyl)amino-3-methoxypyrrole-2-carboxylate

A solution of ethyl 1-methyl-4-amino-3-methoxypyrrole-2-carboxylate (2.75 g, 13.873 mmol) and di-tert-butyldicarbonate (3.37 g, 15.441 mmol) in dioxane (15 mL) was stirred at room temperature under a nitrogen atmosphere for 22 h. The solvent was removed under reduced pressure and the residue was treated with water (100 mL) and the mixture was extracted with CH_2Cl_2 (100 mL). The CH_2Cl_2 layer was washed with water and dried with anhydrous Na_2SO_4, filtered and the solvent was removed on a rotary evaporator to give an orange viscous liquid (4.888 g). ^{1}H NMR of the crude residue was identical with that of an authentic sample.

XII.9: 4-[(tert-Butoxycarbonyl)amino]-1-methyl-3-methoxypyrrole-2-carboxylic acid

Crude ethyl 1-methyl-4-(tert-butoxycarbonyl)amino-3-methoxypyrrole-2-carboxylate (4.88 g, 16.38 mmol) was dissolved in ethanol (16 mL) and a solution of sodium hydroxide, prepared from NaOH (1.376 g) in water (15 mL), was added and the reaction mixture was stirred at

room temperature for six days. Ethanol was removed under reduced pressure. Water (50 mL) was added and the solution was extracted with ether (2 X 50 mL) to remove any unreacted starting ester (0.183 g). The aqueous layer was acidified with H_2SO_4 to pH 2-3 to afford a colorless precipitate. The precipitated product was extracted with ether, dried with anhydrous Na_2SO_4. Filtration and rotary evaporation gave a dirty cream crystalline solid (2.922 g). 4-[(tert-Butoxycarbonyl) amino]-1-methyl-3-methoxypyrrole-2-carboxylic acid prepared by the above new method was identical in all aspects with the authentic sample.

Example XIII - Solution Phase Synthesis of Polyamides

The polyamide solution phase synthesis described below utilizes in part and is a modification and extension of the trichloroacetyl-amine coupling methodology described in: E. Nishiwaki *et al.*, HETEROCYCLES (1988), Vol. 27., 1945-1952

XIII.1 Preparation of O_2N-Py-CO_2TMSE

59.42gm of 2,2,2-trichloro-1-(1-methyl-4-nitro-1H-pyrrol-2-yl)-ethanone (prepared according to Nishiwaki *et al.*, HETEROCYCLES (1988), Vol. 27., 1945-1952) was dissolved in 200 mL of anhydrous THF under an atmosphere of nitrogen. 38.0 mL of 2-trimethylsilylethanol was then added and the solution cooled in an ice bath. 40 mL of 1.0 M lithium-t-butoxide in THF was then added dropwise over a 30-minute period while still cooling with the ice bath. 2.5 mL of acetic acid was then added, and the solvents removed under vacuum. The residue was taken up in a mixture of ether and 5% $NaHCO_3$ and extracted. The ether phase was separated, dried with $MgSO_4$, filtered and the ether removed under vacuum. The product was recrystalized from 85/15 MeOH/H_2O to give 50.39 gm of O_2N-Py-CO_2TMSE. Mass Spectra: m/z+H^+ = 271

XIII.2 Preparation of O_2N-Py-Py-CO_2TMSE

25.05 gm of O_2N-Py-CO_2TMSE and 25.52 gm of ammonium formate were placed under a nitrogen atmosphere and 1.18 gm of palladium on carbon (50% H_2O, 10% Pd dry basis) added. The mixture was refluxed for 90 minutes, cooled, filtered under nitrogen, the solvent removed under vacuum and the residue dissolved in 250 mL of anhydrous acetonitrile under an atmosphere of nitrogen. 27.57 gm of 2,2,2-trichloro-1-(1-methyl-4-nitro-1H-pyrrol-2-yl)-ethanone and 17.5 mL of diisopropylethylamine were added and the mixture refluxed for two hours. The reaction was then cooled and the acetonitrile removed under vacuum. The product was slurried in ice-cold methanol and filtered to give 29.21 gm of O_2N-Py-Py-CO_2TMSE. 2.35 gm of additional product was isolated as a second crop from the mother liquor. Mass Spectra: m/z+H^+ = 393.

XIII.3 Preparation of Im-Py-Py-CO_2TMSE

In a manner similar to the above procedure, 3.00 gm of O_2N-Py-Py-CO_2TMSE was coupled with 2,2,2-trichloro-1-(1-methyl-1H-imidazol-2-yl)-ethanone to give 3.21 gm of Im-Py-Py-CO_2TMSE. Mass Spectra: m/z+H^+ = 472.

XIII.4 Preparation of Im-Py-Py-CO_2H

0.46 gm of Im-Py-Py-CO_2TMSE was dissolved in 4 mL of anhydrous DMF and stirred with 3.0 mL of 1.0 M tetrabutylammonium fluoride in THF under a nitrogen atmosphere for 1.5 hour. The product was purified by flash chromatography using silica gel and 1% acetic acid in 3:1 methylene chloride/THF to give 0.36 gm of Im-Py-Py-CO_2H. Mass Spectra: m/z+H^+ = 372.

XIII.5 Preparation of NO_2-Py-Py-Py-CO_2TMSE

In a manner similar to the procedure describing the preparation of NO_2-Py-Py-CO_2TMSE, 31.51 gm of O_2N-Py-Py-CO_2TMSE was coupled with 2,2,2-trichloro-1-(1-methyl-4-nitro-1H-pyrrol-2-yl)-ethanone to give 37.67 gm of NO_2-Py-Py-Py-CO_2TMSE. Mass Spectra: m/z+H^+ = 515.

XIII.6 Preparation of NH_2-γN(Boc)-Py-Py-Py-CO_2TMSE

In a manner similar to the procedure above, 1 equivalent of O_2N-Py-Py-Py-CO_2TMSE would be coupled with a slight excess of d-1-hydroxybenzotriazolyl 2-[[(1,1-dimethylethoxy)carbonyl]amino]-4-[[(9H-fluoren-9-ylmethoxy)carbonyl]amino]-butanoate to give Fmoc-NH-γN(Boc)-Py-Py-Py-CO_2TMSE. The Fmoc protecting group would then removed by treatment with piperidine to give NH_2-γN(Boc)-Py-Py-Py-CO_2TMSE.

XIII.7 Preparation of Im-Py-Py-γN(Boc)-Py-Py-Py-CO_2TMSE

One equivalent of Im-Py-Py-CO_2H and 1.2 equivalents of O-Benzotriazol-1-yl-*N,N,N',N'*-tetramethyluronium hexafluorophosphate (HBTU) would be dissolved in anhydrous DMF and 1.2 equivalents of diisopropylethylamine added and the solution stirred under a nitrogen atmosphere until the formation of the activated ester is complete. 1 equivalent of NH_2- N(Boc)-Py-Py-Py-CO_2TMSE would then be added and the solution stirred until the amide forming reaction was complete. Diisopropylethylamine may be also be added to aid the second reaction. The product would then be purified, probably by flash chromatography using silica gel and a methylene chloride/THF gradient to give Im-Py-Py-γN(Boc)-Py-Py-Py-CO_2TMSE.

XIII.8 Preparation of Im-Py-Py-γN(Boc)-Py-Py-Py-CO_2H

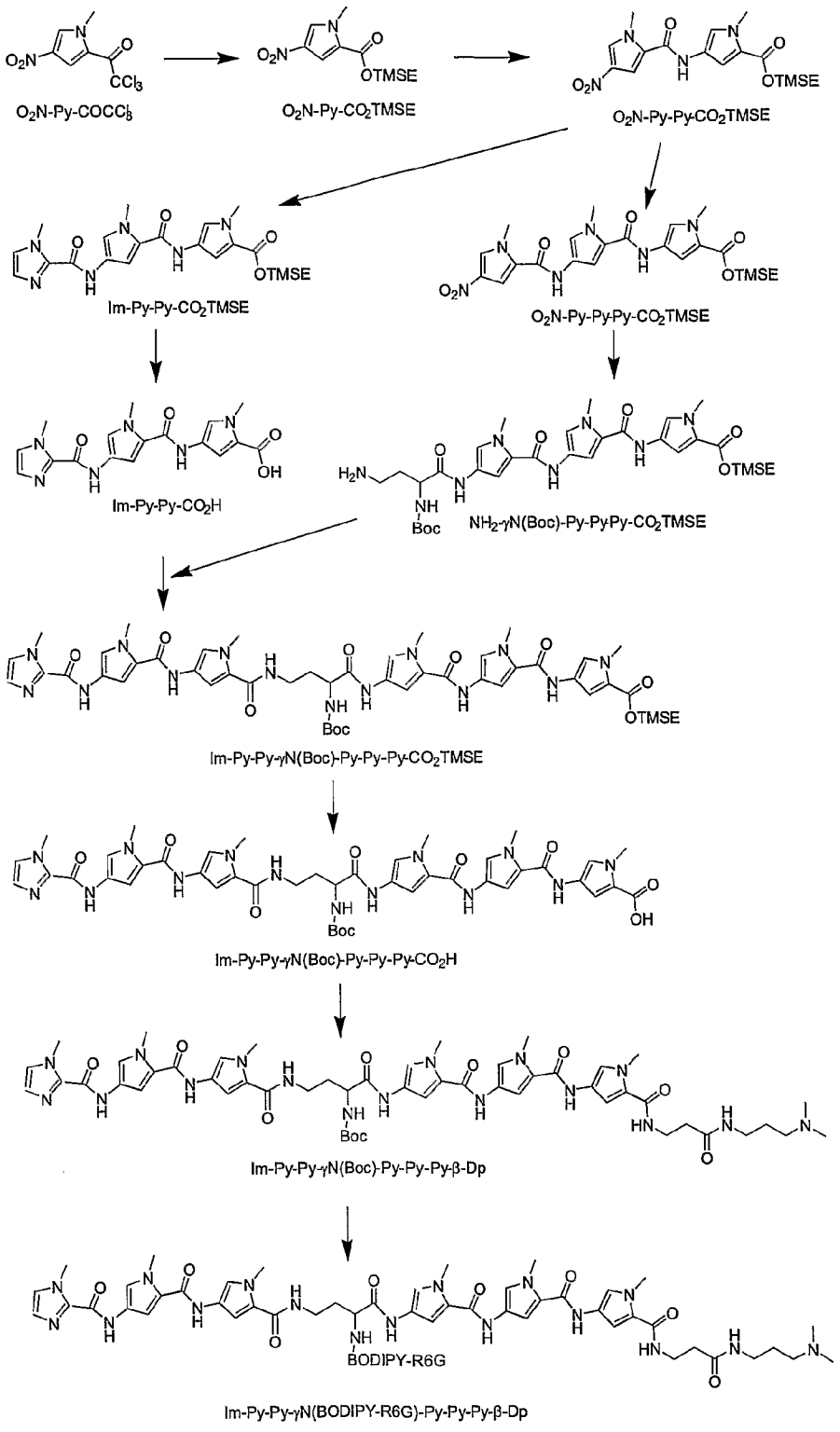
Similar to the method described for the preparation of Im-Py-Py-CO_2H, Im-Py-Py-γN(Boc)-Py-Py-Py-CO_2TMSE would be converted to Im-Py-Py-γN(Boc)-Py-Py-Py-CO_2H.

XIII.9 Preparation of Im-Py-Py-γN(Boc)-Py-Py-Py-β-Dp (prospective example)

Similar to the method described for the preparation of Im-Py-Py-γN(Boc)-Py-Py-Py-CO_2TMSE, Im-Py-Py-γN(Boc)-Py-Py-Py-CO_2H would be coupled with 3-amino-N-[3-(dimethylamino)propyl]-propanamide to prepare Im-Py-Py-γN(Boc)-Py-Py-Py-β-Dp.

XIII.10 Preparation of Im-Py-Py-γN(BODIPY-R6G)-Py-Py-Py-β-Dp (bistrifluoroacetate salt)

One equivalent of Im-Py-Py-γN(Boc)-Py-Py-Py-β-Dp would be dissolved in trifluoroacetic acid (1 mL per 10-50 mg of starting material) and stirred until the Boc group was completely removed from the amine. The trifluoroacetic acid would then be removed under vacuum and the amine dissolved in anhydrous DMF. An excess of diisopropylethylamine would then be added followed by the addition of an anhydrous DMF solution of 4,4-difluoro-5-phenyl-4-bora- 3a,4a-diaza-s-indacene-3-propionic acid, succinimidyl ester (BODIPY® R6G, SE) and the mixture stirred until the reaction was complete. The product would be isolated via reverse phase chromatography, probably using a methanol/water gradient. The product would then be isolated by lyophilization from a t-butanol/water mixture to give the bistrifluoroacetate salt of Im-Py-Py-γN(BODIPY-R6G)-Py-Py-Py-β-Dp.


O2N-Py-COCCl3
O2N-Py-CO2TMSE
O2N-Py-Py-CO2TMSE
Im-Py-Py-CO2TMSE
O2N-Py-Py-Py-CO2TMSE
Im-Py-Py-CO2H
NH2-γN(Boc)-Py-Py-Py-CO2TMSE
Im-Py-Py-γN(Boc)-Py-Py-Py-CO2TMSE
Im-Py-Py-γN(Boc)-Py-Py-Py-CO2H
Im-Py-Py-γN(Boc)-Py-Py-Py-β-Dp
BODIPY-R6G
Im-Py-Py-γN(BODIPY-R6G)-Py-Py-Py-β-Dp

XIII.11 Preparation of Im-Im-Py-Py-γN(Boc)-Py-Py-Py-Py-β-Dp, M1-Boc

Im-Im-Py-Py-γN(Boc)-Py-Py-Py-Py-β-Dp would be prepared using the solution phase methods described for the preparation of Im-Py-Py-γN(Boc)-Py-Py-Py-β-Dp. Alternatively, Im-Im-Py-Py-γN(Boc)-Py-Py-Py-Py-β-Dp would be prepared using the solid phase methods described for the preparation of Im-Py-Py-Py-γ-Im-Py-Py-Py-β-Dp.

XIII.12 Preparation of Im-Im-Py-γN(Boc)-Py-Py-Py-CO_2H, HO_2C-M2-Boc

Im-Py-Py-Nγ(Boc)-Py-Py-Py-CO_2H would be prepared using the methods described for the preparation of Im-Py-Py-γN(Boc)-Py-Py-Py-CO_2H.

XIII.13 Preparation of Im-Im-Py-γN(Boc)-Py-Py-Py-β-CO_2H, HO_2C-β-M2-Boc

Similar to the method described for the preparation of Im-Py-Py-γN(Boc)-Py-Py-Py-CO_2TMSE, Im-Py-Py-γN(Boc)-Py-Py-Py-CO_2H would be coupled with β-alanine, methyl ester to prepare Im-Im-Py-γN(Boc)-Py-Py-Py-β-CO_2Me which would be saponified with sodium hydroxide to prepare Im-Im-Py-γN(Boc)-Py-Py-Py-β-CO_2H. Alternatively, Im-Im-Py-γN(Boc)-Py-Py-Py-β-CO_2H would be prepared using the solid phase methods described for the preparation of Im-Py-Py-Py-γ-Im-Py-Py-Py-β-Dp and cleaved from the solid phase with lithium hydroxide instead of 3-dimethylaminopropylamine.

XIII.14 Preparation of Im-Im-Py-Py-γN(Boc)-Py-Py-Py-Py-β-CO_2H, HO_2C-β-M3-Boc

Im-Im-Py-Py-γN(Boc)-Py-Py-Py-Py-β-CO_2H would be prepared using the methods described for the preparation of Im-Im-Py-γN(Boc)-Py-Py-Py-β-CO_2H.

XIII.15 Preparation of Im-Py-Py-Py-γN(Boc)-Im-Py-Py-Py-β-CO_2H, HO_2C-β-M4-Boc

Im-Py-Py-Py-γN(Boc)-Im-Py-Py-Py-β-CO_2H would be prepared using the methods described for the preparation of Im-Im-Py-γN(Boc)-Py-Py-Py-β-CO_2H.

XIII.16 Preparation of Im-Im-Py-Py-γNH_2-Py-Py-Py-Py-β-Dp, M1

Im-Im-Py-Py-γN(Boc)-Py-Py-Py-Py-β-Dp would be dissolved in trifluoroacetic acid and allowed to stand until the reaction was complete and then the trifluoroacetic acid removed through evaporation to give M1 (Im-Im-Py-Py-γNH_2-Py-Py-Py-Py-β-Dp) as the tetra-trifluoroacetic acid salt.

XIII.17 Preparation of M1-β-M2-Boc

Similar to the method described for the preparation of Im-Py-Py-γN(Boc)-Py-Py-Py-CO_2TMSE, M1 would be coupled with HO_2C-β-M2-Boc to prepare M1-β-M2-Boc.

XIII.18 Preparation of M1-β-M2-β-M3-Boc

Similar to the method described for the preparation of M1-Boc from M1, M1-β-M2-Boc would be converted into M1-β-M2. Similar to the method described for the preparation of M1-β-M2-Boc, M1-β-M2 would be coupled with HO_2C-β-M3-Boc to prepare M1-β-M2-β-M3-Boc.

XIII.19 Preparation of M1-β-M2-β-M3-β-M4-Boc

Similar to the method described for the preparation of M1-Boc from M1, M1-β-M2-β-M3-Boc would be converted

into M1-β-M2-β-M3. Similar to the method described for the preparation of M1-β-M2-Boc, M1-β-M2-β-M3 would be coupled with HO_2C-β-M4-Boc to prepare M1-β-M2-β-M3-β-M4-Boc.

XIII.20 Preparation of M1-β-M2-β-M3-β-M4-BODIPY-FL-X

Similar to the method described for the preparation of M1-Boc from M1, M1-β-M2-β-M3-β-M4-Boc would be converted into M1-β-M2-β-M3-β-M4. Similar to the method described for the preparation of Im-Py-Py-γN(BODIPY-R6G)-Py-Py-Py-β-Dp, M1-β-M2-β-M3-β-M4 would be converted to M1-β-M2-β-M3-β-M4-BODIPY-FL-X.

Example XIV - Preparation of 1-[3-[N-(tert-butoxycarbonyl)amino]propyl]imidazole (It3(Boc))

To a stirred solution of 1-(3-aminopropyl)imidazole (12.518 gm, 100 mmol) in dry acetonitrile (20 mL) was added a solution of di-tert-butyldicarbonate (21.9 gm, 100.34 mmol) in dry acetonitrile (25 mL) over a period of 1 h at room temperature. The reaction mixture was stirred at room temperature under nitrogen for 20 h. The solvent was evaporated in vacuo to give a pale yellow viscous liquid (24.0 gm) that solidified upon standing. High Resolution Mass Spectra: m/z 226.1552 (M+H); Calcd for $C_{11}H_20N3O_2$: 226.1556 (M+H).

Example XV - Preparation of 1-[3-[N-(tert-butoxycarbonyl)amino]propyl]-2-trichloroacetylimidazole (It3(Boc)-COCCl3)

To a stirred solution of trichloroacetyl chloride (4.1 g, 22.55 mmol) in CH_2Cl_2 (35 mL) was added a solution of 1-[3-[(tert-botoxycarbonyl)amino]propyl]imidazole (5.032 g, 22.33 mmol) in CH_2Cl_2 (40 mL) over a period of 2

h at room temperature under a nitrogen atmosphere. The reaction mixture was stirred at room temperature under nitrogen overnight. The reaction mixture was cooled to 0 °C and a solution of triethylamine (3.2 mL, 22.96 mmol) in CH_2Cl_2 (5 mL) was added drop wise over a period of 1 hour. The reaction mixture was warmed to room temperature after 1 hour. The solvent was evaporated in vacuo and the residue chromatographed to give a pale yellow waxy solid that solidified upon standing (3. 38 g, 41%). Mass Spectra: $m/z+H^+$ = 370 with 3 chlorine isotope pattern

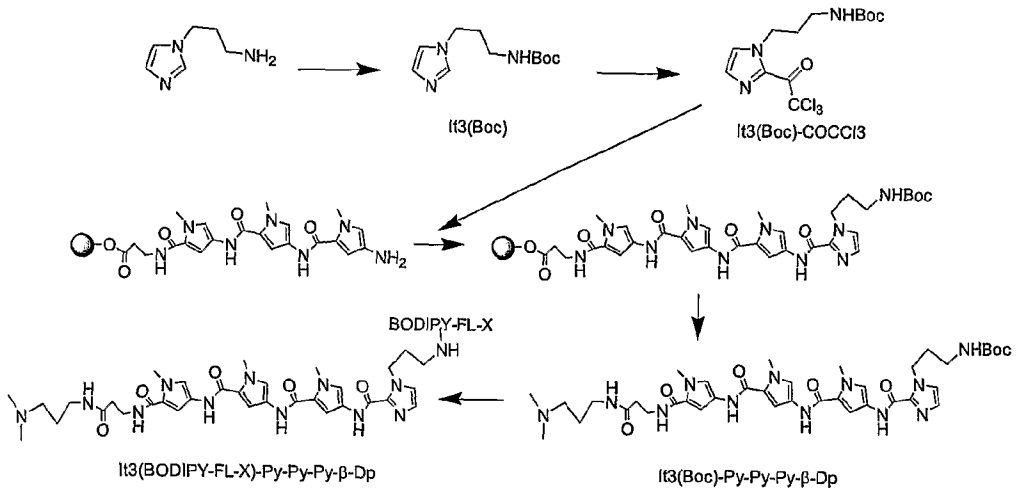
Example XVI - Preparation of It3(Boc)-Py-Py-Py-β-Dp

It3(Boc)-Py-Py-Py-β-Dp was prepared using the solid phase synthesis method described above and using 1-[3-[N-(tert-butoxycarbonyl)amino]propyl]-2-trichloroacetylimidazole for the final coupling step on the solid phase resin. Cleavage with 3-(dimethylamino)propylamine and purification gave It3(Boc)-Py-Py-Py-β-Dp. Mass Spectra: $m/z+H^+$ = 791

Example XVII - Preparation of It3(BODIPY-FL-X)-Py-Py-Py-β-Dp (bistrifluoroacetate salt)

5.83 mg of It3(Boc)-Py-Py-Py-β-Dp was dissolved in 1 mL of trifluoroacetic acid and stirred for 20 minutes. The trifluoroacetic acid was then removed under vacuum and the amine dissolved in 0.5 mL of anhydrous DMF. 50 µL of diisopropylethylamine was then added followed by the addition of 200 µL of a 0.01 M anhydrous DMF solution of 6-((4,4-difluoro-5,7-dimethyl- 4-bora-3a,4a-diaza-s-indacene- 3-propionyl)amino)hexanoic acid, succinimidyl ester (BODIPY® FL-X, SE) and the mixture stirred overnight. The product was isolated via reverse phase

chromatography, using a methanol/water gradient. The product was then isolated by lyophilization from a t-butanol/water mixture to give 2.44 mg of the bistrifluoroacetate salt of It3(BODIPY-FL-X)-Py-Py-Py-β-Dp.



Example XVII: Staining and X,Y-valley depths of X, Y Bull Sperm Nuclei Stained Using Non-Specific Polyamides.

Different oligomeric polyamides are incubated with bovine sperm nuclei in PBS (phosphate buffered saline) and the X,Y-valley split determined using a MoFlo (available from Cytomation, Inc.) FACS ("Fluorescence Activated Cell Sorter") located in the USDA laboratory at Beltsville, MD. An X, Y-valley split of 5% or more is indicative of a clear capability of being separated into at least two populations by flow cytometry. Even in the absence of an X-, Y-valley split, a skilled FACS operator can often or usually select and gate portions of the cells that are to some degree enriched in the

subpopulations occurring in that portion of the scatter plot as illustrated in more detail in Example XVIII below. Runs and results are presented in the following Table.

Table: X, Y-Valley Split of Bull Sperm Nuclei Stained with Polyamides

Run	Polyamide - Dye Conjugate	Stain Conditions	XY Valley Split
1	Im-Py-Py-Py-γ-Im-Py-Py-Py-β - Ta-BO-FL-X	Bull, PBS, pH7.0, 35C, 1h	0%*
2	Im-Py-Py-Py-γ -Im-Py-Py-Py-β - Ta-BO-FL-X	Bull, PBS, pH7.5, 35C, 1h	0%*
3	Im-Py-Py-Py-β-Ta-BO-FL-X	Bull, PBS, pH7.0, 35C, 1h	1-5%
4	Im-Py-Py-Py-β -Ta-BO-FL-X	Bull, PBS, pH7.5, 35C, 1h	10%
5	Im-Py-Py-Py-Py-β-Ta-BO-FL-X	Bull, PBS, pH7.0, 35C, 1h	40%
6	Im-Py-Py-Py-Py-β-Ta-BO-FL-X	Bull, PBS, pH7.0, 35C, 1h	30%
7	Pic-Py-Py-γ-Py-Py-Py-β-Ta-BO-FL-X	Bull, PBS, pH7.0, 35C, 1h	20%
8	Pic-Py-Py-γ-Py-Py-Py-β-Ta-BO-FL-X	Bull, PHS, pH7.5, 35C, 1h	20%
9	Im-Py-γ-Py-Py-β-Ta-BO-FL-X	PBS, pH7.0, 35C, 1h	0%*
10	Im-Py-γ-Py-Py-β -Ta-BOFLX	PBS, pH7.5, 35C, 1h	0%*
11	Im-Im-Py-Py-β-Ta-BO-FL-X	PBS, pH7.0, 35C, 1h	0%*
12	Im-Im-Py-Py-β-Ta-BO-FL-X	PBS, pH7.5, 35C, 1h	0%*
13	Im-Py-β-Py-Py-β-Ta-BO-FL-X	PBS, pH7.0, 35C,	40%

		1h	
14	Im-Py-β-Py-Py-β-Ta-BO-FL-X	PBS, pH7.5, 35C, 1h	40%

* Staining observed

These results and others not shown indicate that non-specific polyamides are capable of use for quantitative separation of X and Y-chromosome bearing genomes using flow cytometry since X-, Y-valley separations of 5% or more are easily separated using flow cytometry. Even where staining is observed but X-, Y-valley separation is not, enrichment of nuclei can usually be accomplished by appropriate gating.

Example XVIII: X,Y Sperm Can Be Separated Using Non-Specific Polyamides

An oligomeric polyamide Im-Py-Py-Py-β-Ta-BO-FL-X capable of permeating cell and nuclear membranes and quantitatively binding to chromosomal DNA was incubated with bovine sperm to evaluate the uptake of polyamide into live bull spermatozoa to evaluate if the dye can quantify DNA content and thereby allow separation of the spermatozoa into X and Y-chromosome bearing populations. The resulting histogram distribution of signal intensities from stained cells indicates that with appropriate gating the sperm can be separated into GES enriched in X chromosome bearing cells or GES enriched in Y chromosome bearing cells.

Semen was collected from a bull using an artificial vagina and extending with tris citric acid buffer containing 1% egg yolk. The sample was sent overnight @ 18°C storage temperature. The suspension was centrifuged (500 x g for 7 minutes) and the supernatant removed. The pellet was re-suspended in TES plus tris buffer (TEST) and the concentration taken. Additional buffer was added to achieve 150 million sperm per mL. A titration experiment was set up using a constant concentration of sperm. Six (6) one mL samples received increasing amounts

(6, 9, 12, 15, 18 and 21 µL) of the polyamide (10 mM stock in DMSO) and was allowed to incubate at 27 °C for 1 hour. A sperm nuclei sample was prepared by ultrasonication (Sonics Vibra) to enable quantitative binding of the dye to the DNA. The sperm nuclei sample was incubated at 35°C for rapid staining of the sperm DNA. A flow cytometer (MoFlow, Cytomation Inc.) was aligned for 488 nm wavelength excitation using FITC coated plastic beads. The laser power was 500mW, and a 515 nm long pass filter was in place to collect spectral emissions above 515 nm. A minimum of 10,000 cells was analyzed flow cytometrically. The sperm nuclei were then processed through the flow cytometer to achieve a splitting index (XY discrimination) of approximately 33% and a coefficient of variation (CV) of 2.24. The peak fluorescent channel for the nuclei was 216. The results of the live sperm are shown in the table below.

Stain condition	Motility Pre-flow	Splitting index	CV	Peak fluorescent channel
Nuclei	NA	33%	2.24	216
6 µL dye	70%	0	3.35	213
9 µL dye	40%	0	2.53	221
12 µL dye	70%	0	2.80	218
15 µL dye	70%	0	2.78	217
18 µL dye	70%	0	2.81	214
21 µL dye	70%	0	3.03	212

The frequency histograms showed one population of cells and the peak fluorescence was similar in value to the nuclei indicating that the fluorescence signal was similar between the live and dead cells. Therefore the uptake of polyamide dye conjugate was maximal for the live cells. The CV values were near to the CV of the nuclei indicating a near saturation of the sperm DNA of the live cells. Clumping of the cells (sperm heads

binding to one another) was evident in all of the live samples and this may have influenced the resolution of the X and Y populations in the live samples. The data indicate that the sperm were saturated with the dye and if one were to use a gating procedure, then significant XY separation could be achieved resulting in subpopulations enriched in either the x-bearing or Y-bearing chromosomes. The gating procedure might, for example, collect populations of the 33%-most intense cells and the 33% least intense cells. It is common practice with separation techniques such as chromatography that, in those cases where baseline peak-to-peak separation is not achieved, one can take fractions and assess their purity by a second analysis. Those skilled in the art will appreciated that significant separation can occur prior to the appearance of a distinct valley. Thus, so long as the CV's approach those for which a valley will occur, which can be simulated to an extent by using stained nuclei for setting up the FACS, an preliminary analysis of a sperm sample will give a range of fluorescence intensities for that sample, as indicated by the channel numbers in which counts appear. This range can then be used to establish a gating procedure for a production run. Cells in, for example, the lower 1/3 of the intensity range and cells in the upper 1/3 of the intensity range would be collected to give populations that are enriched in Y and X sperm, respectively. Then, the populations isolated by this method can be assessed by staining and sorting with Hoechst 33342 by FACS as is known or by FISH or by or other techniques known in the art.

Example XIX: Separation of X,Y Sperm Using Specific Polyamides.

A polyamide would be designed and constructed using a sequence selected from the DNA sequence referred to in Example VII (See, e.g., McGraw et al., Nucleic Acids Research 10389 (1988)), tested for Y-chromosome specificity and affinity and used for flow cytometric separation of porcine sperm into GES enriched in X chromosome bearing cells and GES enriched in Y chromosome bearing cells.

Example XX: Separation of X, Y Sperm Using Selected Combinatorially Produced Polyamides.

A combinatorial library of polyamides would be generated, e.g., using the methodology of Boger et al., "Total Synthesis of Distamycin A and 2640 Analogues: A solution-phase combinatorial approach to the discovery of new, bioactive DNA binding agents and development of a rapid, high-throughput screen for determining relative DNA binding affinity or DNA binding sequence selectivity," J. Am. Chem. Soc.122, 6382 (2000), screened against X and Y chromosomes isolated as described in Section 6, and polyamide(s) selectively binding to the X or Y chromosomes selected and used for flow-cytometric separation of bovine or porcine sperm into GES enriched in X chromosome bearing cells and GES enriched in Y chromosome bearing cells.

Example XXI: Separation of X, Y Sperm Using FRET System Polyamide-Dye Conjugates.

Example XIII would be repeated using FRET system polyamide dye conjugates and GES produced enriched in X chromosome bearing cells and GES enriched in Y chromosome bearing cells.

While the invention has been described in terms of various working, preferred and prospective embodiments, the invention in its various aspects is not restricted thereto but by the claims hereafter appended as interpreted in accordance with applicable law.

WHAT IS CLAIMED IS:

1. A method for producing subpopulations of reproductive cell genomic DNA comprising:

a. providing a population of genomic DNA that has been contacted under conditions effective for binding with a DNA stain comprising an oligomeric polyamide effective for binding target DNA sequences in the genomic DNA, the oligomeric polyamide optionally conjugated to a dye; and

b. based on resulting staining of genomic DNA producing a subpopulation of genomic DNA enriched or impoverished in target DNA sequences.

2. The method of claim 1 wherein the population of reproductive cell genomic DNA is selected from the group consisting of nuclei of reproductive cells, nonviable gametes, viable non-motile gametes and viable motile gametes.

3. The method of claim 1 wherein the population of reproductive cell genomic DNA is obtained from the group of reproductive cells comprising gametes and their immediate predecessor cells giving rise to gametes by cell division.

4. The method of claim 1 wherein the population of reproductive cell genomic DNA is obtained from the group of reproductive cells consisting of sperm, oocytes and ova of a selected species of agriculturally important animal.

5. The method of claim 1 wherein the target DNA sequence is selected to have an abundance distribution within the population of genomic DNA permitting the separation of genomic DNA into at least two subpopulations of genomic DNA based on a difference in relative abundance of the target DNA sequence in genomic DNA allocated to respective subpopulations.

6. The method of claim 1 wherein the target DNA sequence is selected from extragenic sequences of the genomic DNA.

7. The method of claim 1 wherein the target DNA sequence is selected from genic sequences of the genomic DNA.

8. The method of claim 1 wherein the target DNA sequence is randomly distributed in the genome.

9. The method of claim 1 wherein the target DNA sequence is non-randomly distributed in the genome and occurs preferentially on one of the sex chromosomes.

10. The method of claim 1 wherein the target DNA sequence is a Y-chromosome specific extragenic repeating sequence.

11. A composition of matter comprising reproductive cell genomes stained with target DNA specific DNA stain comprising an oligomeric polyamide optionally conjugated with a dye.

12. The composition of claim 11 wherein the reproductive cells are selected from the group consisting of sperm,

ova and oocytes and immediate predecessor cells from which the cells are derived by cell division.

13. The composition of claim 11 wherein the reproductive cells are selected from the group consisting of ovine, porcine, bovine, and equine sperm.

14. The composition of claim 11 wherein the reproductive cells comprise living sperm selected from the group consisting of bovine and porcine sperm.

15. The composition of claim 11 wherein the target DNA sequence is selected from extragenic sequences contained in the reproductive cell genome.

16. The composition of claim 11 wherein the target DNA sequence is selected from genic sequences contained in the reproductive cell genome.

17. The composition of claim 11 wherein the oligomeric polyamide dye conjugate comprises a single oligomeric molecular structure for preferentially binding the target DNA sequence.

18. The composition of claim 11 wherein the oligomeric polyamide comprises a mixture of at least two oligomeric molecular structures for adjacently binding the target DNA sequence.

19. The composition of claim 11 wherein the reproductive cell genomes are contained in living bovine sperm.

20. The composition of claim 11 wherein the reproductive cell genomes are contained in living porcine sperm.

21. A method for separating reproductive cell genomes contained in living sperm comprising:
    a. staining the genomes with a stain selected from the group consisting of DNA target specific oligomeric polyamides and polyamide dye conjugates capable of quantitative binding to the genomes for enabling the separation of X- and Y-chromosome bearing sperm; and
    b. producing a subpopulation of living sperm enriched in proportion of one of X- and Y-bearing sperm.

22. The method of claim 21 wherein step b. comprises use of a fluorescence activated cell sorter.

23. The method of claim 21 wherein step a. comprises use of DNA target specific polyamides of less than 8 heterocyclic monomer units.

24. The method of Claim 21 wherein step a. comprises use of DNA target specific polyamide or polyamide dye conjugates preferentially binding extragenic repeat sequences of the Y-chromosome.

25. The method of claim 21 wherein the sperm comprise bovine sperm.

26. The method of claim 21 wherein the sperm comprise porcine sperm.

27. The method of claim 21 wherein the method further comprises providing the resulting subpopulation of sperm for artificial insemination.

28. The method of Claim 21 wherein the method further comprised providing the resulting subpopulation of sperm for in vitro fertilization.

29. The method of claim 21 wherein the oligomeric polyamide dye conjugate comprises a single oligomeric molecular structure for preferentially binding adjacent the target DNA sequence.

30. The method of claim 21 wherein the oligomeric polyamide comprises a mixture of at least two oligomeric molecular structures for adjacently binding the target DNA sequence.

31. A method of isolating selected chromosomes of reproductive cells comprising:

a. staining a chromosome in a mixture of chromosomes using a fluorescent oligomeric polyamide or polyamide dye conjugate preferentially binding to a target DNA sequence in the selected chromosome; and

b. producing a subpopulation of chromosomes enriched in the mixture of chromosomes by fluorescence activated cell sorting responsive to the fluorescent oligomeric polyamide.

32. The method of Claim 31 wherein the reproductive cells are sperm selected from bovine, ovine, porcine and equine sperm.

33. The method of Claim 31 wherein the reproductive cells are sperm and the Y chromosome is preferentially stained using an oligomeric polyamide preferentially binding to a repeating extragenic sequence of the Y chromosome.

34. A composition of matter comprising mammalian embryos stained with a polyamide dye conjugate specific for target DNA sequences therein.

35. The composition of Claim 34 wherein the mammalian embryos are selected from embryos of agriculturally significant species.

36. The composition of Claim 34 wherein the embryos comprise embryos in the range from zygotes through and including the morula stage of development.

37. The composition of Claim 34 wherein the polyamide dye conjugate is selective for extragenic repeat sequences.

38. The composition of Claim 34 wherein the polyamide dye conjugate is effective for distinguishing between male and female embryos.

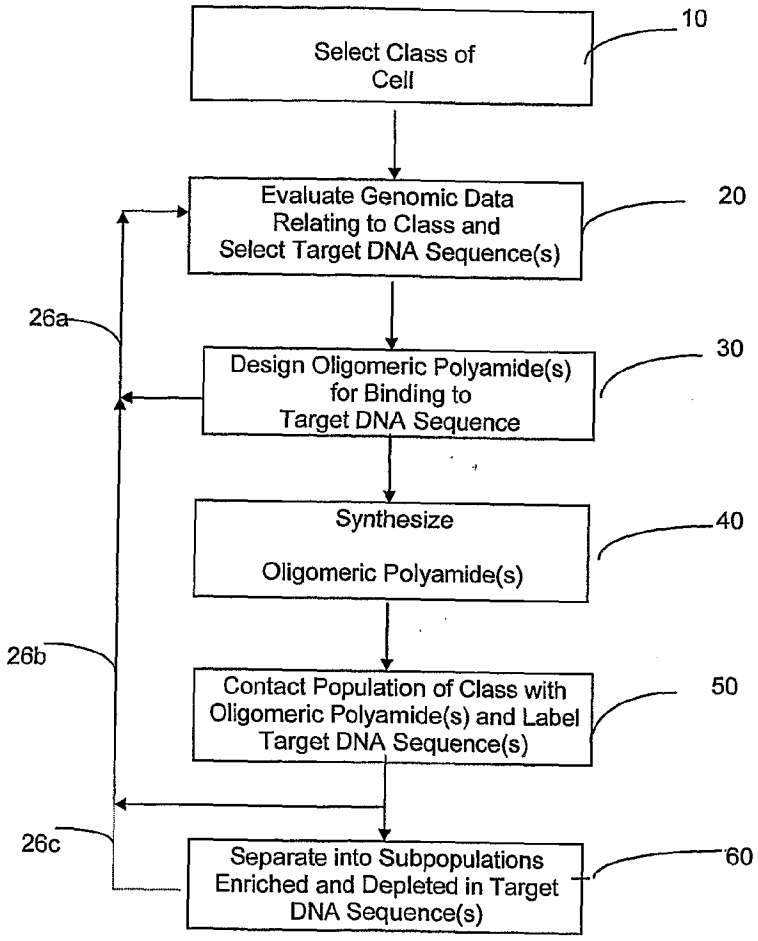

Select Class of Cell
10
Evaluate Genomic Data Relating to Class and Select Target DNA Sequence(s)
20
26a
Design Oligomeric Polyamide(s) for Binding to Target DNA Sequence
30
Synthesize Oligomeric Polyamide(s)
40
26b
Contact Population of Class with Oligomeric Polyamide(s) and Label Target DNA Sequence(s)
50
26c
Separate into Subpopulations Enriched and Depleted in Target DNA Sequence(s)
60

FIGURE 1

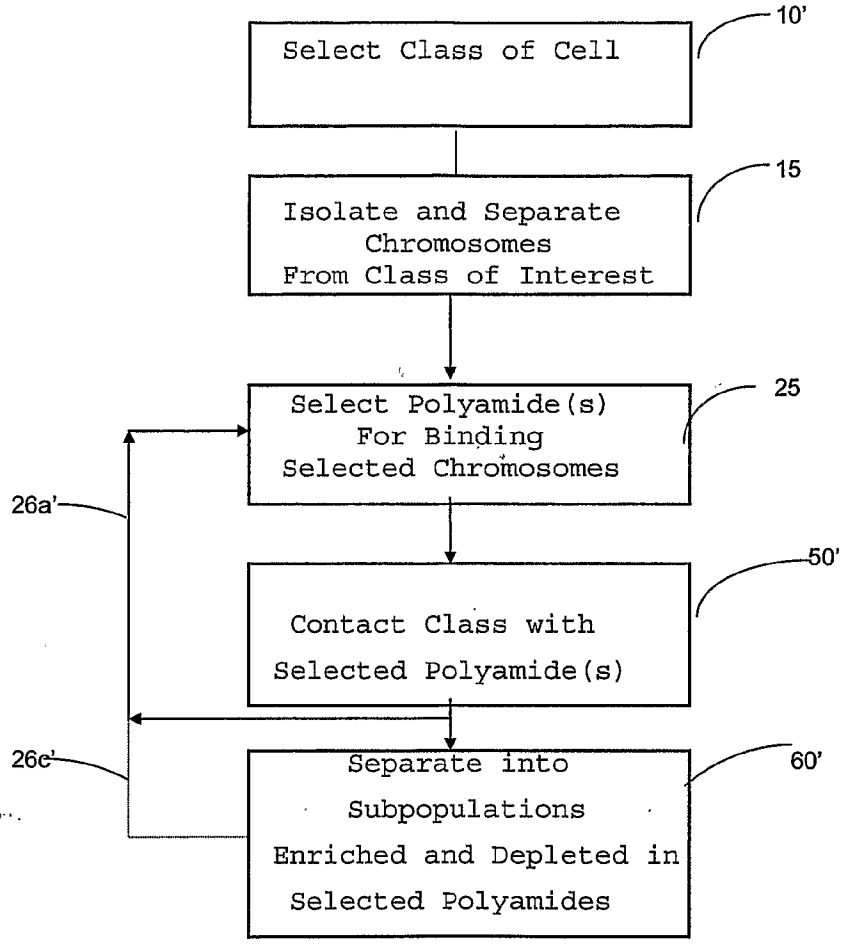

Select Class of Cell
10'
Isolate and Separate Chromosomes From Class of Interest
15
Select Polyamide(s) For Binding Selected Chromosomes
25
26a'
Contact Class with Selected Polyamide(s)
50'
26c'
Separate into Subpopulations Enriched and Depleted in Selected Polyamides
60'

FIGURE 1A

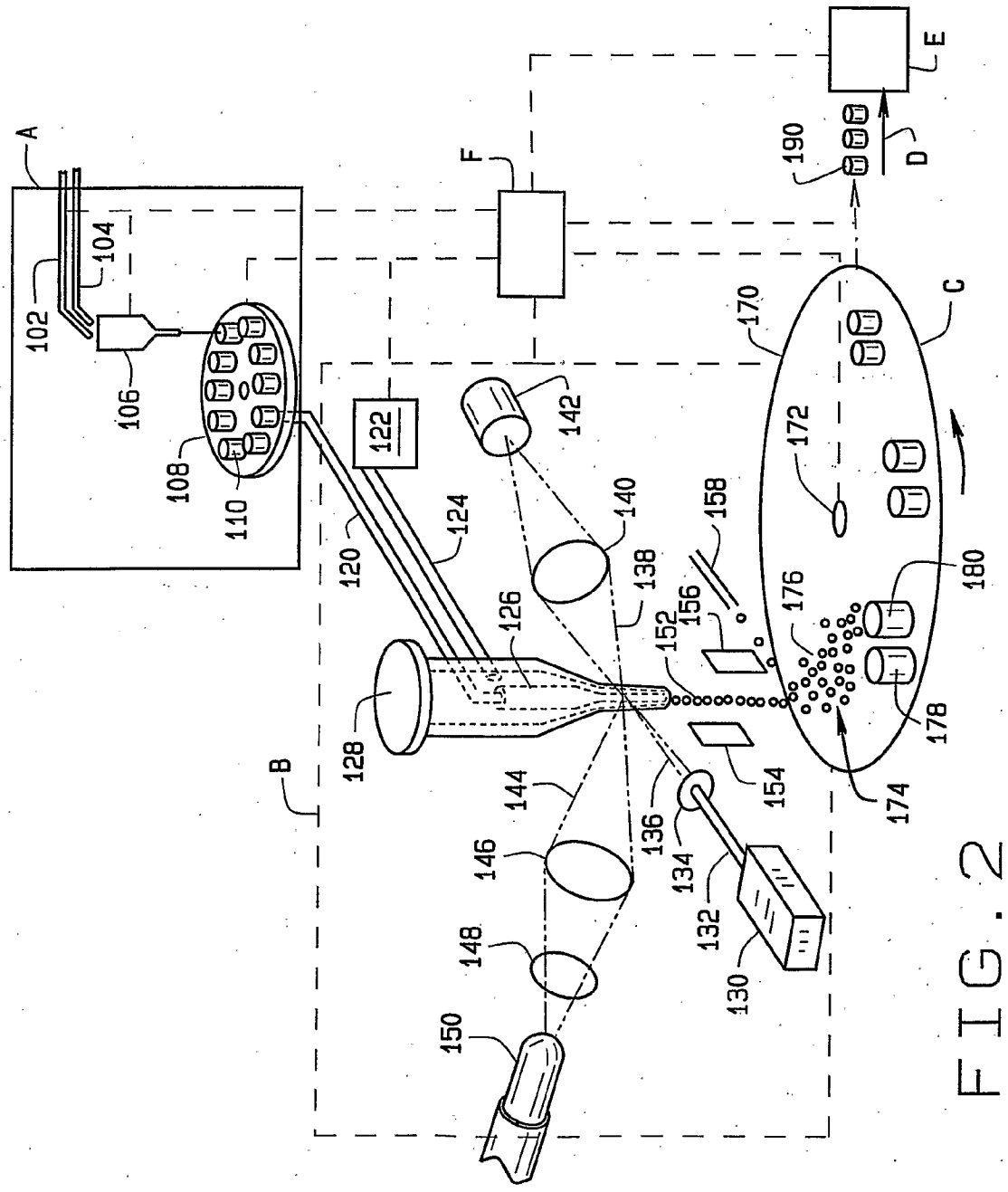

A
102
104
106
108
110
120
122
124
126
128
B
140
142
144
146
148
150
136
134
132
130
138
152
154
156
158
F
E
190
D
C
170
172
174
176
178
180

FIG. 2

SEQUENCE LISTING

<110> Pharmacia Corporation
Didion, Bradley A.
Bashkin, James K
Woodard, Scott S.

<120> STAINING AND SORTING GENOMES AND CHROMOSOMES USING SEQUENCE-SPECIFIC POLYAMIDES

<130> 37-21(51737)C

<150> US 60/316,729
<151> 2001-09-01

<160> 9

<170> PatentIn version 3.1

<210> 1
<211> 11
<212> DNA
<213> Unknown

<220>
<223> oligonucleotide

<400> 1
tatgtattta t 11

<210> 2
<211> 11
<212> DNA
<213> Unknown

<220>
<223> oligonucleotide

<400> 2
tatgtatcta t 11

<210> 3
<211> 11
<212> DNA
<213> Unknown

<220>
<223> oligonucleotide

<400> 3
tatgtatgta t 11

<210> 4
<211> 10
<212> DNA
<213> Unknown

<220>
<223> oligonucleotide

<400> 4
gtagtattat 10

<210> 5
<211> 10
<212> DNA
<213> Unknown

<220>
<223> Oligonucleotide

<400> 5
gtagtcttat 10

<210> 6
<211> 18
<212> DNA
<213> Unknown

<220>
<223> Oligonucleotide

<400> 6
gtagtattat agtattat 18

<210> 7
<211> 6
<212> DNA
<213> Unknown

<220>
<223> Oligonucleotide

<220>
<221> misc_feature
<222> (1)..(1)
<223> Nucleotide in position 1 is unknown.

<220>
<221> misc_feature
<222> (4)..(4)
<223> Nucleotide in position 4 is unknown and may be present one or mor
e times.

<220>
<221> misc_feature
<222> (6)..(6)
<223> Nucleotide in position 6 is unknown.

<400> 7
nwgnwn 6

<210> 8
<211> 26
<212> DNA
<213> Bos taurus

<400> 8
atgcggtcct gggcaattgg caggct 26

<210> 9
<211> 23
<212> DNA
<213> Bos taurus

<400> 9
tggaaaaggt aaggaaaagt tct 23